THE UNIVERSITY
of EDINBURGH

The Sheep Lung Microbiota

By Laura Glendinning

Thesis Submitted for the Qualification Doctor of Philosophy, The University of Edinburgh, 2017

Abstract

Until recently it was assumed that the healthy mammalian lung did not harbour a microbiota, unlike other body sites. However, through the use of sequencing based technologies this has been shown to not be the case. Low biomass communities of microbes can be identified in the healthy lung and the lung microbiota in various diseases states has been shown to differ from these 'healthy' communities. The sheep respiratory microbiota is of interest from both an animal health perspective and due to the potential use of the sheep as a large animal model for studying the lung microbiota. In this thesis I seek to characterise the composition and variability of the sheep lung microbiota; the differences between the sheep upper and lower respiratory tract bacterial communities and to assess whether exhaled breath condensate collection can be used as a non-invasive lung microbiota sampling method. To study the bacterial communities present in samples I have used 16S rRNA gene sequencing and analysis.

In Chapter 3 I examine the inter-individual and spatial variability present within the sheep lung microbiota. Protected specimen brushings were collected from three lung segments in six animals at three time-points. In a separate sheep a greater number of brushings was taken (n=16) in order to examine the amount of variability over a smaller spatial scale. I find that there can be large differences between the bacterial communities isolated from different locations within the lung, even over short distances. Samples also cluster by the sheep from which they were taken, indicating a host specific influence on the lung microbiota. In Chapter 4 I compare whole lung washes and oropharyngeal swabs from 40 lambs in order to examine the differences between the upper and lower respiratory tract microbiotas. I find that oropharyngeal swabs separate into rumen-like or upper respiratory tract-like bacterial communities. Despite the fact that in humans the upper and lower respiratory microbiotas have been shown to have similar compositions, the sheep lung microbiota samples in this study do not resemble either oropharyngeal samples or reagent only controls.

In my first two results chapters, lung sampling methods were used which involved either anaesthesia combined with a bronchoscopic procedure (Chapter 3) or samples being taken from dead animals (Chapter 4). In Chapter 5 I assess whether there is a less invasive way of taking lung microbiota samples from a living individual, both to minimise the procedural stress on animals used as models and to increase the pool of potential volunteers for human lung microbiota studies. I compared samples taken via protected specimen brushings to samples taken via exhaled breath condensate collection, a less invasive sampling technique. I find that condensate samples contain less bacterial DNA and different bacteria than brushing samples, indicating that it is unlikely they could be used as a replacement for invasive sampling methods. In my final results chapter I compare the results across Chapters 3, 4 and 5 to identify bacteria which occur consistently in the sheep lung and could therefore potentially be described as core lung microbiota members.

In conclusion, while I have found that there are large differences between the sheep lung microbiota and that which has previously been described in humans, the sheep can still be of use as a model in studies where these differences would not have a significant impact, such as in Chapter 5 of this thesis. I have identified several bacterial members of the core sheep lung microbiota which in future it would be interesting to better characterise and to assess whether they play a role in sheep health.

Lay summary

During health, communities of microorganisms can be found inhabiting the mammalian lung; these are commonly referred to as the 'lung microbiota'. The study of the lung microbiota is a relatively new field as until recently it was assumed that the healthy lung was sterile. However, using DNA sequencing this has been shown to be incorrect. In this thesis I examine the types of bacteria which can be found in the sheep lung. This area of research is of interest as not only is the sheep an important agricultural animal but it is often used as a model in studies of respiratory diseases due to the similarity of the sheep respiratory and immune systems to those of humans. In order to work out what types of bacteria are in respiratory samples, I first extract all of the DNA from the samples then from this DNA mixture I sequence a gene known as the 16S rRNA gene. This gene is found in all bacteria and its DNA sequence varies depending on the species of bacteria it originates from. Therefore, if we examine the 16S rRNA genes in a sample we can work out which types of bacteria it originally contained.

In my first results chapter I look at how lung bacterial communities differ at different lung locations and between sheep. I find that there is often a lot of variability between samples taken from different parts of the same lung, even when these samples were taken a few centimetres away from each other. However, I also find that despite this variation, lung microbiota samples from an individual sheep are more similar to each other than they are to samples from another sheep. This might mean that the conditions in an individual sheep's lungs help determine what kind of bacteria inhabit it.

In my second results chapter I compare throat swabs and lung fluids from a larger number of animals (40 lambs) in order to see whether the types of bacteria in the upper respiratory tract are different from those in the lower respiratory tract. I find that, unlike in humans, there are large differences between the lower and upper sheep respiratory tract microbiotas. In this results chapter I sampled the lung microbiota by collecting lung fluid from dead animals while in my first results chapter lung samples were taken using invasive bronchoscopic procedures in live animals. I wanted to see whether there was a less invasive way of sampling the lung microbiota from live individuals, as reducing the invasiveness of sampling could minimise the procedural stress on animals which are part of lung microbiota studies and could also increase the pool of potential human volunteers for these studies. Therefore, in my third results chapter I compare samples taken via bronchial brushings and exhaled breath condensate collection, a sampling technique which does not require bronchoscopy. I find that far less bacterial DNA was able to be collected from exhaled breath condensate than from bronchial brushings and the bacterial communities found in the two sample types were different. This means that it is unlikely that exhaled breath condensate collection will be able to be used to replace more invasive lung microbiota sampling techniques. In my final results chapter I compare the results from all of my previous results chapters to identify bacteria which can often be found in the sheep lung.

I conclude that while there are differences between the lung microbiota in sheep and that which had previously been found in humans, the sheep can still be useful as a large animal model in some lung microbiota studies. I also conclude that in future studies it would be interesting to study some of the bacteria I identified as common inhabitants of the sheep lung, to attempt to understand whether they play a role in sheep health.

Author's declaration

I declare that this thesis was composed by myself and that the work contained therein (including publications) is my own, except where explicitly stated throughout the thesis and within this declaration. This work has not been submitted for any other degree or professional qualification. Contributions from other individuals towards the work contained in this thesis include:

Chapter 3: This chapter was published as 'Variability of the Sheep Lung Microbiota' in Applied and Environmental Microbiology (1). Anaesthesia was induced and monitored by Steven Wright, David Collie and Peter Tennant. Bronchoscopic procedures were carried out by Peter Tennant. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording.

Chapter 4: This chapter was submitted to the journal Microbiome on 19/07/2016 as a research paper. Lung fluid samples were collected by myself with the help of Steven Wright and David Collie. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording.

Chapter 5: This research was published as 'Microbiota in Exhaled Breath Condensate and the Lung' in Applied and Environmental Microbiology (2). Colistimethate sodium delivery and EBC collection from conscious animals were performed by myself and Gerry McLachlan. Anaesthesia was induced and monitored by Steven Wright, David Collie and Peter Tennant. EBC collection in anaesthetised animals was performed by myself, David Collie and Peter Tennant. Bronchoscopic procedures were carried out by David Collie and Peter Tennant. Blood was collected post mortem by Steven Wright, David Collie and Gerry McLachlan. HPLC was performed by Andy Gill. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording.

Acknowledgements

Firstly I'd like to thank all of the members of my group, particularly Gerry, David, Steven and Liz, for all the support they've given me over my PhD and the patience they have shown when I have stuck their heads on pictures of sheep during my acknowledgement slides in talks. I'd also like to thank all of my family for their general support, particularly my mam and dad who helped me develop my love of science and very rarely showed that they were fed up with me asking so many questions all the time. A special thank you also to Mr Moody, my secondary school biology teacher, without whom I wouldn't have chosen to be a microbiologist.

Jolinda Pollock needs thanking for so many things it isn't really practical for me to list them all here. We've written papers together, helped each other out with frustrating mothur based problems…but what I'd say I valued most is just having someone to occasionally rant with (Why do people still not include negative controls!!!).

One thing I have learned during my PhD is that large animals are ridiculously expensive and without donations of samples from Kenny Rutherford, Cheryl Ashworth and Emily Clarke I wouldn't have been able to do anywhere near as many experiments as I did. A massive thank you also to all of the people who worked on the farm, particularly Peter Tennant who always managed to fit my experiments into his incredibly busy schedule. I'd like to thank Bob Fleming and Andy Gill for helping me with a bunch of technical, methodological stuff and my thesis committee for their encouragement and support. I'd also like to thank all of the reviewers/editors of my papers for their advice and guidance and for occasionally giving me a bit of a giggle (the giraffe comment was fantastic).

Thanks to the BBSRC, not only for funding my project and giving me lots of chances to develop my extra-academic skills, but also for paying me a decent enough salary so that I could not only pay for food, bills etc. (which was touch and go during my masters) but could also greatly increase my board/card gaming collection over the course of my PhD.

Talking about board gaming, I would also like to thank my geeky board gaming and roleplaying friends for providing a nice fantasy world to escape into when I just couldn't stare at the computer any longer without going insane. Hopefully Cleva will one day learn that not all situations can be solved by buying people drinks and Folami will eventually tell you what gender they are (or maybe they won't).

Finally, I'd like to thank Julien for being generally awesome.

Publications

1) Glendinning L, Free A. Supra-organismal interactions in the human intestine. Front Cell Infect Microbiol. 2014 (Volume 4):47.

2) Glendinning L, Nausch N, Free A, Taylor DW, Mutapi F. The microbiota and helminths: Sharing the same niche in the human host. Parasitology. 2014 (Volume 141):1255-71.

3) McGorum BC, Pirie RS, Glendinning L, McLachlan G, Metcalf JS, Banack SA et al. Grazing livestock are exposed to terrestrial cyanobacteria. Vet Res. 2015 (Volume 46):16.

4) Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P et al. Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate sodium. PLoS ONE. 2015 (Volume 10):e0142097. **Appendix 1**.

5) Glendinning L, Wright S, Pollock J, Tennant P, Collie D, McLachlan G. Variability of the sheep lung microbiota. Appl Environ Microbiol. 2016 (Volume 82):3225-38.

6) Glendinning L, Wright S, Tennant P, Gill AC, Collie D, McLachlan G. 2017. Microbiota in exhaled breath condensate and the lung. Appl Environ Microbiol. 2017 (Volume 83): e00515-17

Contents

Abbreviations

AMOVA: Analysis of Molecular Variance

BAL: Bronchoalveolar Lavage

CMS: Colistimethate Sodium

COPD: Chronic Obstructive Pulmonary Disease

EBC: Exhaled Breath Condensate

FUT2: Fucosyltransferase 2

HPLC: High Performance Liquid Chromatography

IgA: Immunoglobulin A

IU: International Units

LCD: Left Caudal Diaphragmatic

LVD: Left Ventral Diaphragmatic

OTU: Operational Taxonomic Unit

PBS: Phosphate Buffered Saline

PCOA: Principle Coordinate Analysis

PCR: Polymerase Chain Reaction

PSB: Protected Specimen Brushing

qPCR: Quantitative Polymerase Chain Reaction

RA: Right Apical

RC: Right Cardiac

RCD: Right Caudal Diaphragmatic

RVD: Right Ventral Diaphragmatic

SD: Standard Deviation

SEM: Standard Error of the Mean

SRA: Sequence Read Archive

UV: Ultra Violet

Chapter 1: Introduction

"We cannot fathom the marvellous complexity of an organic being; but on the hypothesis here advanced this complexity is much increased. Each living creature must be looked at as a microcosm--a little universe, formed of a host of self-propagating organisms, inconceivably minute and as numerous as the stars in heaven." - Charles Darwin (3)

The microbiota can be defined as the commensal, symbiotic and pathogenic microbes which inhabit a particular ecological niche. Microbiota have been associated with even the most simple of animal life, such as nematodes (4), indicating that the microbiota is likely to have evolved alongside animals from an early evolutionary stage. This co-evolution has allowed animals and the microbes that inhabit them to benefit nutritionally and has also helped prevent the colonisation of the host by pathogenic organisms and contributed towards the development of the adaptive immune system (5-8). The microbiota has also been found to play a role in various diseases states including inflammatory, autoimmune and infectious diseases, allergies, cancer and psychiatric disorders (9, 10). Sequencing the 16S rRNA genes contained within a sample allows the taxonomic identification of the bacteria present without being biased by whether they are culturable. Using this technique, great strides have been made in our understanding of the microbiotas at different body sites and the impact they have in health and disease. However, our knowledge of the lung microbiota is still in its infancy.

In this thesis I will explore the composition, diversity and dynamics of the lung microbiota in sheep. There is a pressing need for a large animal model of the lung microbiota, as the lungs of rodents and other small mammals are anatomically very different from the lungs of humans and lung sampling techniques are often quite invasive which may deter human volunteers from taking part in studies. As well as being important as livestock animals, sheep are commonly used as models of respiratory disease due to the anatomical and immunological similarity of their lungs to those of humans (11, 12). In the following introductory chapter I will review the current literature on the lung microbiota, mainly focusing on human studies, and I will then describe how sheep may have a place in this area of research. Lastly, I will describe the specific methodological problems which arise during lung microbiota studies.

1.1 The lung microbiota

1.1.1 Classical view of lung sterility

Until recently, it was widely believed that the healthy human lung was a sterile environment (13, 14). This view was supported by the belief that the lung would be rapidly cleared of any potential bacterial colonisers. The lung environment provides a suitable habitat for the growth of many potentially harmful microorganisms due to the availability of nutrients, the presence of oxygen and a temperature of ~37°C. It is therefore essential that a robust immune response to potential pathogens is present in the lower airways. Microbes attempting to colonise the lungs will first encounter a thick layer of mucus which coats the epithelium of the upper airways and the bronchi (15). Mucus containing bacteria is then transported up the airways via the action of ciliated epithelial cells whereupon it is either swallowed or expelled by coughing or sneezing. The bronchioles and alveoli do not contain a mucus layer but several antibacterial molecules are secreted into the lung lumen at these sites as well as in the upper airways.

Polymeric IgA molecules make up the majority of antibodies present in the lung mucosa but IgM and to a lesser extent IgG and IgE can also be found at the lung epithelium (16). Antimicrobial products are also secreted into the airway surface liquid, including lysozymes, lactoferrins, cathelicidins, defensins and various other peptides (17). Another layer of protection for the lung epithelium is added by the presence of alveolar macrophages which are able to phagocytose invading microbes and initiate and resolve inflammatory responses.

It was previously assumed that the immune mechanisms described above were sufficient to keep the healthy lung sterile. However, several papers published in 2010 and 2011 questioned this idea, describing the presence of bacteria in a number of healthy humans using 16S rRNA gene sequencing (18-20). Since these publications, various studies have been performed to attempt to better characterise the mammalian lung microbiota.

1.1.2 Establishment of the lung microbiota and the maintenance of homeostasis

One of the first papers to examine the healthy lung microbiota concluded that the upper and lower respiratory tracts were indistinguishable from one another, except that a far lower bacterial biomass could be found in the lungs in comparison to oral washes and oropharyngeal swabs (20). A potential conclusion from this study would be that the human lung does not contain a distinct microbiota and the presence of bacteria is merely due to their passive diffusion from further up the respiratory tract. This is one of several postulated models of how the lung microbiota may be established, more of which will be discussed below.

Microaspiration is normal in healthy adults and studies in humans point to this being the origin of the majority of the microbes inhabiting the lung (21). The neutral dispersal model hypothesises that once microbes have dispersed from the upper respiratory tract into the lung there is no selective pressure from the lung environment for particular bacterial species and any changes in species abundance are merely due to ecological drift. As a result, the lung microbial communities would be highly similar to those of the aspirated fluid. If this were the case then one would expect to see not only the same bacteria in the upper and lower respiratory tracts but also the same bacteria at different locations within the lung, as regional differences in pH, temperature and oxygen level (producing 'microhabitats' (22)) should not affect the lung microbiota composition. Venkataraman et al. concluded that the neutral model accurately described what they observed when comparing oral washes and bronchoalveolar lavage (BAL) fluid in 62 healthy volunteers (23). While Venkataraman et al. did not find any bacterial operational taxonomic units (OTUs) which appeared to be consistently under environmental pressure from the lungs, other studies have identified OTUs which were significantly disproportionally represented in the lungs, including both bacteria (24) and fungi (25).

Another model which has been suggested to explain how microbes migrate to the lung and how these microbial communities achieve homeostasis is the adapted island model of biogeography. This model is an adaptation of the equilibrium model of island biogeography proposed by MacArthur and Wilson in 1963 (26). The original model postulates that the species richness of an island is due to the balance of immigration from the main landmass and extinction of species on the island. Islands which are smaller and farther from the land mass would have low species richness whereas large islands located near the land mass would have high species richness (**Fig. 1.1**). The adapted island model hypothesises that, in a similar fashion to communities on islands closer to the mainland, the lung microbial communities located nearer the upper respiratory tract would be richer than those which were located in the lower airways (27). This difference in richness was demonstrated in a study of 15 healthy humans where communities in the right upper lung lobe were found to be richer and to more closely resemble the upper respiratory tract than communities from more distal lobes (28).

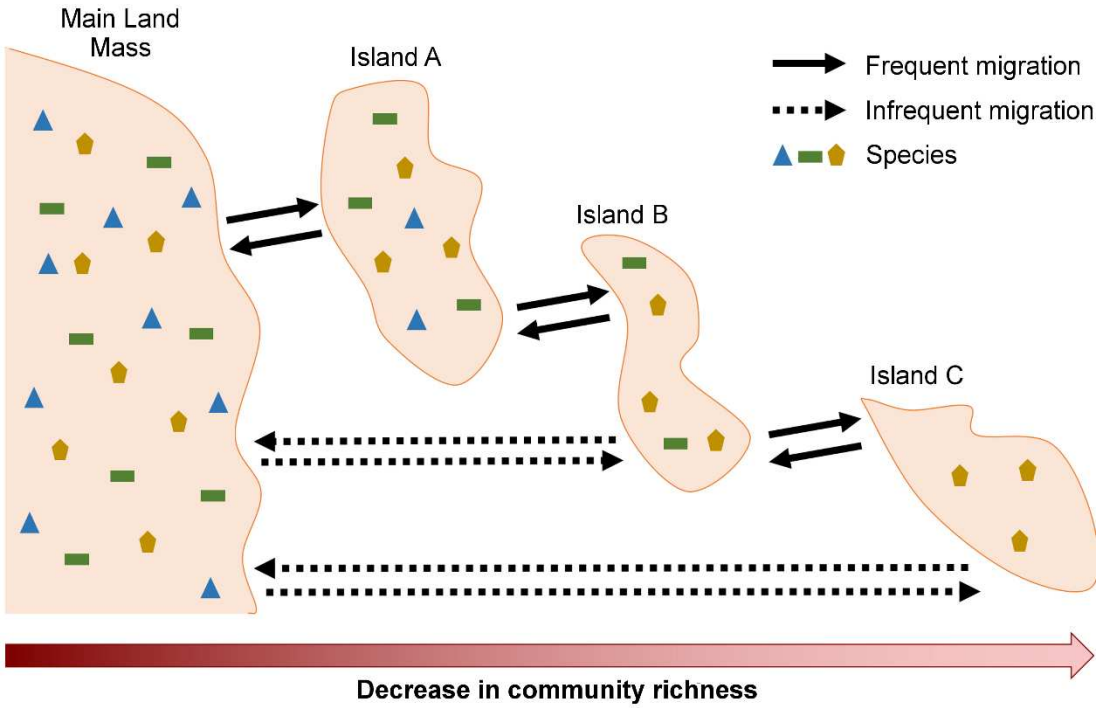


Figure 1.1: A basic diagram depicting the island biogeography model. This model states that community richness is established by a combination of immigration and elimination elements. In this simple example there are three islands (originally containing no species and of approximately equal size) which are located different distances from the mainland (originally containing numerous species). The extinction rate of these species is assumed to be the same on each island thereby community richness is related only to immigration from the mainland or other islands. Islands which are further away from the source community (the mainland) have a lower community richness than those which are nearer. When applied to the lungs, this model would assume that the further locations in the lung were from the source community (the upper respiratory tract), then the less rich they would be (if elimination rates are the same across the lung).

For all microbial communities there are three main factors which drive their composition: immigration, elimination and reproduction. While microaspiration of oral bacteria is the main source of immigrant species in the human lung (29), it is probable that environmental bacteria contained in air are also inhaled in smaller quantities. It is also possible that bacteria could migrate from other body sites if they entered the bloodstream, as occurs in sepsis-induced acute respiratory distress syndrome where the lung microbiota is temporarily enriched with gut microbes (30). The elimination of lung bacteria is due to a combination of mucociliary clearance, coughing and immune mechanisms, while reproduction is believed to have only a small impact on the healthy human lung microbiota.

Another factor which may drive the composition of the lung microbiota is early life microbial exposure. While no well-controlled studies have been published on the establishment and development of the lung microbiota in healthy human neonates, this has been examined in mice (31). The quantity of bacteria in the lungs was found to significantly increase over the first two weeks of life and communities shifted from being predominantly composed of Gammaproteobacteria and Firmicutes towards communities which were dominated by Bacteroidetes. This study did not examine the origin of these bacteria but it is probable that they originated from the upper respiratory tract, as in adults. Babies born by caesarean section are more likely to develop respiratory disease (32, 33) which has led researchers to examine whether different modes of delivery lead to different microbes inhabiting the upper respiratory tract. One study examined several time-points over the first months of life and observed a similar development of the nasopharyngeal communities to the microbiota at other body sites, with keystone microbes colonising the site, followed by the development of a more complex, niche specific microbiota over time (34). This study also showed that there were some differences in the microbiota compositions based on mode of delivery. Differences in the upper respiratory tract communities of babies delivered either naturally or by caesarean section have also been observed in other studies (35, 36). Whether these changes in the upper respiratory microbiota correlate with changes in the lung is currently unknown.

The proportion of bacterial DNA detected in healthy human lungs originating from dead vs live bacteria is also currently unknown. Pezzulo et al. concluded that 94% of the bacterial 16S rRNA gene sequences found in pig lung homogenates and 63% of bacteria in pig BAL fluid originates from dead bacteria (37). It has also been shown that removing the DNA originating from dead bacteria in cystic fibrosis lungs leads to different communities of bacteria being identified by sequencing (38). It must be noted that it is not required that bacteria be alive for them to be able to elucidate an immune response, meaning that dead bacteria may still influence the development and maturation of the immune system and therefore have an impact on respiratory disease.

1.1.3 The lung microbiota in health

As the main origin of lung microbes during health appears to be the upper respiratory tract, it is worth briefly considering the composition of the microbial communities present at these sites. While the lower airways have historically been considered sterile, the upper airways including the oral and nasal cavities, pharynx and larynx have long been known to harbour a diverse community of commensal organisms. It is more accurate to think of the upper respiratory tract as multiple niches harbouring diverse microbial communities rather than one unified microbiota. In a study by Bassis et al. the human lung microbiota was found to more closely resemble oral wash samples than nasal swabs (39), therefore I will focus my discussion on this area of the upper respiratory tract.

In humans the predominant bacterial phyla present in the oral microbiota are Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria with smaller numbers of Spirochaetes, Fusobacteria, TM7, Synergistetes, SR1, Tenericutes and Cyanobacteria (40). This niche also contains small numbers of fungi, archaebacteria, protozoa and viruses (41). Studies examining non-human mammals have found large differences between humans and other mammals in terms of their upper respiratory tract microbiota compositions (42-46). When using animal models it is therefore important to keep in mind these differences. Some of the common oral microbiota found in sheep will be covered more specifically in Section 1.2.

The composition of the human oral microbiota is more dependent on environment than genetics. Twins have been shown to have more similar salivary microbiotas to each other than when compared with the general population yet monozygotic twins show equally similar salivary microbiotas to dizygotic twins (47). Also, dysbiosis of the oral microbiota is well known to cause periodontal disease which is closely related to environmental factors (48).

Unlike the upper respiratory tract where 65.6% of the microbial taxa identified by sequencing have also been cultured (40), the microbial communities in the lungs have been far less amenable to culturing, therefore the most common microbial taxa found in the lungs were only recently identified. In a study of healthy human smokers and non-smokers, the taxonomic groups *Haemophilus* and Enterobacteriaceae were found in higher abundance in the lung in comparison to the upper airways and the opportunistic pathogen *Tropheryma whipplei* was found in a quarter of individuals (24). While it is likely that in humans the majority of lung microbes are sourced from the upper airways, this does not necessarily mean that the relative abundances of microbes will be the same in both niches as certain microbes may be able to survive/multiply within the lung more easily or may be aspirated more often.

It also does not necessarily follow that every individual will have a lung microbiota which is composed of microbes from the upper airway. Two papers by Segal et al. identified two lung bacterial pneumotypes (49, 50). One pneumotype reflected saline background controls while the other pneumotype was similar to supraglottic samples. This indicates that while some individuals do aspirate upper airway bacteria into their lungs, this is not observed at a detectable level in others. These studies also highlighted a link between the presence of supraglottic bacteria in the lung and increased inflammation.

Once bacteria have entered the lungs, it is also important to ask which lung structures they colonise. Yun et al. used fluorescence *in situ* hybridisation to attempt to visualise bacteria in the mouse lung (51). In germ-free and specific pathogen free (SPF) mice no fluorescence signal was discovered, whereas in non-SPF mice and wild mice fluorescence was detected around the epithelial lining of the alveoli, despite no obvious signs of inflammation being present. The researchers stated that while some of the fluorescence signals seemed to indicate single bacteria, there are also appeared to be

biofilm-like structures present. Dickson et al. also suggested that members of the lung microbiota may form biofilms, as removing the host cells from BAL fluid reduces the quantity of bacteria and changes the types of bacteria identified (52). This suggests that a significant proportion of the lung bacteria are host cell associated which may arise through biofilm formation, bacteria being located intracellularly in macrophages or by specific cell to cell adhesion mechanisms.

Regardless of how and where bacteria are located in the lung during health, they must come into contact regularly with the host immune system. That the presence of supraglottic bacteria is linked to an inflammatory phenotype points to a relationship between the lung microbiota and the immune system. Changes in mouse lung bacterial communities over the first two weeks of life have been correlated with decreased house dust mite aeroallergen responsiveness (31). The same paper found that this decrease in responsiveness did not occur in germ-free mice over this two week period and concluded that airway microbiota formation induces regulatory immune cells in early life. However, several papers have suggested that the gut microbiota is also linked to changes in lung immunity (53, 54). It is therefore not possible to conclude whether the difference between germ-free and wild type mice in this study are due to the absence of the lung or gut microbiota. It has also been suggested that studies which show that the gut microbiota affects susceptibility to respiratory infections may unintentionally be changing the lung microbiota and that this may be the true cause of changes in susceptibility (55). However, as of yet, this is purely speculative.

It must also be stated that the vast majority of lung microbiota research in humans has been carried out in individuals from developed countries and that demographic information for study participants is often not reported. A study comparing healthy adult Malawians who were exposed to high or low levels of airborne particulate matter due to domestic combustion of biomass fuels found significant changes in the abundances of several lung bacteria (56). This may indicate that environment can have an effect on the lung microbiota during health and would suggest that lung microbiota research should be carried out in more diverse human populations who are exposed to different environmental conditions.

1.1.4 The lung microbiota in disease

Some non-infectious diseases were known to be correlated with bacterial lung infection before the common use of 16S rRNA gene sequencing, including cystic fibrosis and chronic obstructive pulmonary disease. Modern sequencing techniques have been able to further explore the dynamics of lung microbiota changes during these diseases and have also discovered relationships between the lung microbiota and various other diseases. I will briefly describe some of these relationships in order to give context as to why the study of the lung microbiota is clinically important.

Cystic Fibrosis

Due to a combination of culture-based and sequencing-based studies, cystic fibrosis has probably the best understood lung microbiota phenotype of any disease. Lung inflammation occurs in cystic fibrosis patients shortly after birth but bacteria are not always cultured when inflammation starts to be detectable (57). This could be due to the presence of pathogens in unsampled sections of the lung; infection being at low and thereby undetectable levels; unculturable bacteria being the cause of infection or dead bacterial components being the original initiators of the inflammatory response. As it is possible to detect bacterial DNA in the lungs of young children with cystic fibrosis, even at the age of one month (58, 59), a lack of cultured bacteria is unlikely to indicate a lack of lung bacteria altogether.

During childhood *Streptococcus*, *Haemophilus*, *Pseudomonas*, *Staphylococcus* and *Achromobacter* are the most dominant genera identified by 16S rRNA gene sequencing, whereas in adults *Pseudomonas*, *Burkholderia*, *Streptococcus*, *Haemophilus* and *Staphylococcus* are dominant (60). This reflects culture based studies where *Staphylococcus aureus* and *Haemophilus influenzae* are commonly cultured from infant patients and as individuals age the quantities of these bacterial species in the lungs decreases and *Pseudomonas aeruginosa* becomes the principal infecting bacterium (61). *P.aeruginosa* is antagonistic to specific members of the lung microbiota such as *S.aureus* (62, 63), potentially leading to its dominance at later disease stages. One interesting finding is that stable cystic fibrosis patients, who are not demonstrating a decline in lung function, have more inter-connected, ecologically-dependent communities of lung microbes than patients demonstrating a decline in lung function, suggesting that their lung microbiotas may be more resistant to the effects of opportunistic pathogens such as *Pseudomonas* (64). This suggests the possibility of deliberately colonising the lung with certain bacterial species to form a more resilient community of microbes. However, the field of lung microbiota research is still in its infancy and it is likely to be some time before such a therapy would be feasible.

The main factors which increase the ability of bacterial species to colonise the cystic fibrosis lung are related to changes in the airway surface liquid. One theory states that the airway surface liquid in cystic fibrosis lungs contains high levels of salt which have been shown to inactivate antimicrobial defensins and that water levels are also depleted producing a more viscous airway surface liquid and impaired mucus clearance (65-67). Another theory states that while the cystic fibrosis airways do not contain high salt levels, the cystic fibrosis airway surface liquid has a reduced pH and thereby has decreased antimicrobial peptide activity (68). These factors may increase the ability of bacteria to colonise this environment but they do not explain why specific bacterial species are more able to do so. The ability of *H.influenzae* and *P.aeruginosa* to infect the cystic fibrosis lung may be partly due to their ability to form biofilms on the airway epithelia (69, 70). Indeed, the highly viscous cystic fibrosis lung mucus may promote biofilm formation in *P.aeruginosa* colonies due to restricted bacterial motility and small molecule diffusion (71).

Fodor et al. and Cuthbertson et al. found that once a community of bacteria had established itself in the adult cystic fibrosis lung, the overall community structure remained relatively stable over time, even during exacerbations and after antibiotic treatment (72, 73). This seems to agree with Li et al. who found that after treating exacerbations with antibiotics the lung microbiota was only transiently changed during the period of treatment (74) and Whelan et al. who found that exacerbations were not consistently associated with changes in the microbiota (75). Although the adult lung microbiota appears stable over time, this does not mean that there are not differences based upon where in the lung the bacteria are located. Spatial variability in the bacterial lung microbiota and in *P.aeruginosa* clones has been observed (76, 77). This variability may be one of the reasons why there are differences in the bacterial communities extracted from lung microbiota samples taken from cystic fibrosis patients using different methods (78, 79).

COPD (chronic obstructive pulmonary disease)

After cystic fibrosis, the next most studied disease in relation to the lung microbiota is COPD. The term COPD describes a collection of lung diseases caused by long-term exposure to irritants such as tobacco or cooking fire smoke, causing lung inflammation leading to narrowing of the airways and reduced air flow in the lungs. COPD patients are frequently subject to respiratory infections which may be partially due to impaired mucus clearance (80).

The effect of COPD on lung microbiota diversity is as yet not well agreed upon. Patients with moderate and severe COPD were shown by Erb-Downward *et al.* to have reduced lung microbiota diversity in comparison to healthy controls and to display, as in cystic fibrosis, spatial variability in their lung microbiota (19). However, Pragman et al. observed that rather than showing a decrease in diversity, patients with moderate COPD actually had more diverse lung bacterial communities than controls and patients with severe disease displayed an even greater increase in diversity (81). This agrees with a previous finding by Sze et al. (82). The sample sizes in these studies were quite small, particularly for the severe COPD groups which consisted of <10 individuals. A more recent study of healthy smokers (n=8), non-smokers (n=11) and individuals with COPD (n=18) used both culture and DNA sequencing based methods to analyse the lung microbiota (83). This study concluded that microbiota diversity is decreased in COPD lungs in comparison to the other groups. To confuse matters further, a study comparing patients with mild COPD to healthy controls found no difference in diversity between the two groups (84).

Despite disagreements about diversity, all of the studies mentioned above do agree that there are differences in terms of the microbiota composition of COPD affected vs healthy lungs. The mechanisms which drive airway inflammation in COPD are not very well understood but it has been suggested that the local microbiota may be partially responsible. Richmond et al. tested this hypothesis in a strain of mice which did not secrete IgA, as reduced IgA secretion has been previously

shown in COPD affected lungs (85). These mice developed a lung phenotype which closely resembled COPD and changes were observed in their lung microbiota compositions. However the same mouse strain, raised in germ-free conditions did not develop airway inflammation, leading the authors to speculate that IgA deficiency may lead to the innate immune systems always being activated by members of the lung microbiota which could lead to airway remodelling.

Other diseases

As well as cystic fibrosis and COPD, several other diseases have been correlated with changes in the lung microbiota. Studies examining these diseases can suggest how lung microbiota homeostasis is established and maintained during health.

For example, lung transplantation studies can provide us with a clue as to how the lung microbiota is established, as at the point of transplantation the lung should contain microbes from only the original host but it will be exposed to different types of bacteria from the upper respiratory tract in the new host. In lung transplants of cystic fibrosis patients, the strains of *P.aeruginosa* found in patients' lungs pre and post-transplantation are usually found to be the same, whereas the general community structure of the microbiota is not as consistently similar (86). Bernasconi et al. suggested that the presence of different types of bacterial communities in transplanted lungs is correlated with changes in the lung inflammatory markers, specifically lungs which were dominated by one phylum (dysbiosis) showed different immune cell activation profiles than lungs which were not dominated by one phylum (no dysbiosis) (87).However, I feel that the use of phylum abundance as a measurement of dysbiosis in this study is overly simplistic as it is possible to have a highly diverse ecosystem which contains many species which originate from the same phylum, therefore caution should be taken when interpreting these findings.

It is well known that the immune system plays a key role in maintaining microbiota homeostasis at other body sites, such as the gut, and that the absence of a microbiota leads to changes in immune system development (88, 89). Preventing members of the microbiota from crossing epithelial barriers decreases the immune system's exposure to these commensal microbes and thereby prevents an excessive inflammatory response. Immunoglobulin A (IgA) molecules specific to members of the gut microbiota are secreted into the gut lumen whereupon they bind to these bacteria and prevent them from crossing the epithelial barrier (90). T_{reg} cells also play an important role in maintaining gut microbiota homeostasis; when T_{reg} cells are depleted in the host this leads to an increase in $CD4^+$ T_H cells which are specific to commensal bacteria and this further leads to increased inflammation (91). The presence of different types of α-defensins has also been linked to changes in the gut microbiota composition of mice (92).

The microbiota is also important in the development of the immune system, with germ free mice developing significantly smaller gut lymphoid follicles than specific pathogen free mice (93). Specific commensal gut bacteria have also been linked to the proliferation of particular T-cell subsets. For example, segmented filamentous bacteria have been shown to increase Th17 cell differentiation in the gut (94). It has also been demonstrated that neomycin-sensitive members of the microbiota play a role in regulating the mucosal immune response to viral pathogens in the lung by inflammasome activation (95).

While the interactions between the lung microbiota and the immune system are not currently as well understood as those in the gut, the fact that changes can be observed in the lung microbiota in patients with immunological disease indicates that the immune system is likely to be a key driver of lung microbiota composition. That HIV infection leads to changes in the lung microbiota (25), particularly during advanced disease (96), adds more evidence that the immune system is instrumental in driving microbial population dynamics. Fungi have also been found to be more likely to infect the lungs of those with severe, uncontrolled asthma, indicating a potential role of airway inflammation (97). A study examining lipopolysaccharide induced lung injury in mice adds further evidence that inflammation plays a role in changing the lung microbiota, and specifically attributes this to the outgrowth of opportunistic pathogens from within the host (98). Lipopolysaccharide induced lung injury resulted in an increase in endogenous Proteobacteria in the lung which the authors suggested was due to increased substrates in the lung environment which these bacteria could use for growth. The authors also highlighted that the types of bacteria seen in their lung injury model were similar to those observed during complications in ventilated patients in intensive care units. It has also been suggested that disruption of the lung microbial communities (eg. by inflammation or antibiotic treatment) may lead to changes in immune function, producing a more inflammatory immune environment and thereby further driving the colonisation of the lung by pathogens which are able to cause chronic infection (99).

The lung microbiota of ventilated human patients has recently been examined in two studies via the collection of endotracheal aspirate and it was found that bacterial diversity decreased as time on the ventilator increased, suggesting the outgrowth of a small number of bacterial species (100, 101). The authors of the 2016 study suggest that if it were possible for sequence based methods to be used in clinical practice it might aid clinicians, as during their study it may have allowed for earlier clinical intervention or for antibiotic treatment to have been tailored when unexpected species of bacteria were found to be dominating the lung.

One study in non–cystic fibrosis bronchiectasis patients also highlighted the importance of genetics in determining the composition of the lung microbiota (102). The *FUT2* (fucosyltransferase 2) gene encodes for a α(1, 2)fucosyltransferase which facilitates the expression of ABH, Lewis^{b} and Lewis^{y} glycans on mucosal surfaces. As some microbes have been found to utilise these glycans to adhere to mucosal surfaces (103) this group hypothesised that individuals who had homozygous loss-

of-function mutations in the *FUT2* gene may have different airway microbiota compositions to those who had a functional copy of the gene. This study found that individuals who did not carry a functional copy of the *FUT2* gene were found to have lung microbiota profiles which were significantly less likely to be dominated by *Pseudomonas aeruginosa* than those patients who had at least one functional copy of the gene. It is well known that the gut microbiota composition is influenced by a wide range of both environmental and genetic factors (104) and it seems logical to assume that this would also be the case for the lung microbiota, although the impact of genetics on the healthy human lung microbiota has not currently been studied. It is also possible that sequencing could be used to predict whether a patient was likely to experience future disease exacerbations, as differences in lung microbiota compositions have been associated with worse future clinical outcomes in some diseases, such as in non–cystic fibrosis bronchiectasis (105). Specifically, the presence of *P.aeruginosa* (which is linked to a worse clinical outcome) is able to competitively exclude *H.influenzae***,** which when present is related to a less severe disease outcome (106). Interestingly, unlike in bronchiectasis, *Haemophilus* colonisation has been correlated to a poorer response to inhaled fluticasone treatment in asthma (107). Differences in the resident microbiota of sheep have also been shown to affect their chances of developing lung disease and I shall discuss this further in the next section.

1.2 The sheep lung

1.2.1 Use as a model of human disease

Sheep lungs (**Fig. 1.2**) are often used as a model in respiratory disease research due to their increased similarity to human lungs in comparison to rodent models, including having similar immunological responses to allergic stimuli and structural similarity to human lungs (11, 12, 108). Despite the fact that lamb lungs at birth are more advanced than in humans and that they demonstrate less morphological changes postnatally, they are still more useful as a model of human lung development than rodents due to the far less advanced state of rodent lungs at birth (109).

Sheep have been used as models for asthma (110), transfusion-related lung damage (111), pulmonary embolism (112), acute respiratory distress syndrome (113) and gene therapy for cystic fibrosis (114). My research group developed a model of *P.aeruginosa* infection in the sheep lung which was used to investigate how *P.aeruginosa* infection and treatment with intravenous antibiotics changed the lung microbiota in sheep (115) (**Appendix 1**). Briefly, we found that bacterial DNA could be identified in the lungs of *Pseudomonas* infected and uninfected animals. Despite finding that lung microbiota samples from within the same animal were more similar to each other than to samples taken from a different animal, we observed high levels of inter- and intra-individual variation in the bacterial communities within the lungs which led us to ask the research questions discussed during Chapter 3. We also observed a decrease in the relative abundance of Gram negative bacteria (excluding

Pseudomonadales) after sheep were treated with intravenous colistimethate sodium (CMS), indicating that it may be possible to cause changes in the sheep lung microbiota through targeted antimicrobial treatment. This finding led us to explore whether aerosolised CMS treatment could lead to larger changes in the lung microbiota due to increased deposition of the drug in the airway lining fluid (see Chapter 5) (116).

Another large animal which is commonly used in respiratory research is the pig (108), and it could be argued that as pigs do not ruminate they may be a better model for studying the lung microbiota. However, there are several advantages to using sheep over pigs as respiratory models. While absent in pigs, both sheep and humans have networks of small airway connections which are present between adjacent lung segments, allowing for collateral ventilation (12, 108). Currently it is unknown whether this would cause microbial communities to be dispersed throughout the lung in a different manner in animals which were not capable of collateral ventilation. Also, sheep are generally more docile than pigs and are easier to handle, allowing for greater ease when conducting experiments in conscious animals (such as is performed in Chapter 5 of this thesis).

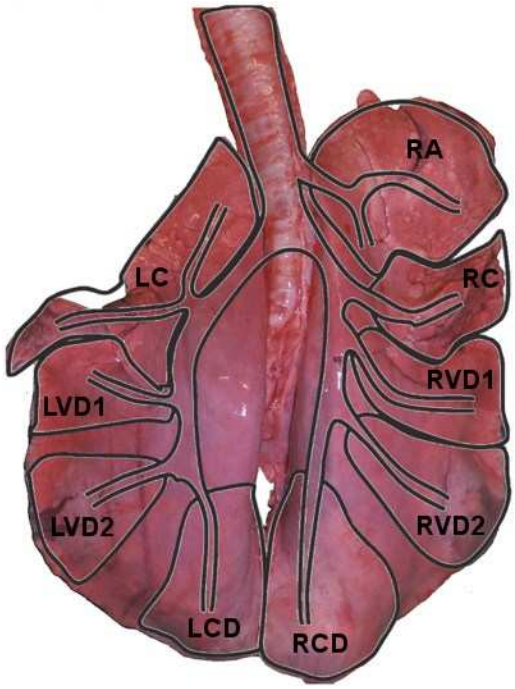


Figure 1.2: Diagram of the sheep lung divided into segments. These include the right apical (RA), right cardiac (RC), right ventral diaphragmatic 1 and 2 (RVD1 and RVD2), right caudal diaphragmatic (RCD), left caudal diaphragmatic (LCD), left ventral diaphragmatic 1 and 2 (LVD1 and LVD2) and the left cardiac (LC) segments. Adapted from Collie et al. (117).

1.2.2 The sheep lung microbiota during health

Prior to the study by my group summarised in the previous section, no one had previously examined the healthy lung microbiota of sheep using non-culture based methods. A small number of culture based studies had been performed and showed variable results.

One study examining the differences between normal and pneumonic sheep lungs found that bacteria could be isolated and grown from 9% of healthy sheep lungs in comparison to 66% of pneumonic lungs (118). The most common microbe isolated from both lung types was *Pasteurella haemolytica* (aka *Mannheimia haemolytica*), a common member of the upper respiratory tract microbiota and a causative agent of the respiratory disease pasteurellosis in sheep. In pneumonic lungs, *Moraxella catarrhalis* (a common member of the upper respiratory tract microbiota but found very rarely in healthy lungs) was unlikely to be found by itself but was commonly found co-infecting with *P.haemolytica*. It was also observed that in pneumonic sheep lungs if *P.haemolytica* or *M.catarrhalis* were present it was very rare for them to not also be present in the nasal cavity. While we cannot be certain that the presence of these bacteria in the lungs is due to outgrowth from the upper respiratory tract, it is reasonable to assume that, as in humans, the sheep lung microbiota would be established by microaspiration of microbes form the upper respiratory tract. This study also highlighted that *P.haemolytica* and *M.catarrhalis* were likely to dominate the nasal flora during the autumn, potentially indicating that the sheep respiratory microbiota may be seasonal.

A separate study examining the respiratory tracts of apparently healthy sheep at an Ethiopian abattoir also reported seasonal changes in the respiratory microbiota, showing a substantially larger ratio of Gram positive to Gram negative bacteria in February to April but a slightly larger Gram negative to Gram positive ratio in November to January (119). This study identified far higher colonisation rates of the 'healthy' lung than the previously discussed study, with 62.5%.of lungs showing bacterial colonisation (with the nasal cavity, tonsils and trachea being colonised 97.9%, 93.8% and 79.2% of the time respectively). Several microbes were cultured from the lungs of over 10% of the animals tested including *Arcanobacterium pyogenes* (now named *Trueperella pyogenes*), *Bacillus* spp., *Citrobacter* spp., coagulase-negative staphylococci, *Corynebacterium* spp., *Enterococcus faecalis*, *Klebsiella* spp., *P.haemolytica*, *Micrococcus* spp., *Pasteurella multocida*, *Pseudomonas* spp. and *Enterococcus* spp.

In contrast to these results, which seem to indicate that colonisation of the healthy lung is quite common, a study of Nadji and Somalian sheep found that lung colonisation was very rare and the only bacterium reported to have been found in healthy animal lungs was *Corynebacterium pseudotuberculosis*, which was found in 8.7% of the Najdi sheep (120). This differing level of lung colonisation across studies could be due to many factors including the breed of sheep and environmental conditions.

1.2.3 The upper respiratory tract as a harbour for disease causing microbes

Various microorganisms have been linked to lung disease in sheep including *P.haemolytica*, *Pasteurella* spp., *Mycoplasma ovipneumoniae*, *Chlamydia psittaci*, *Mycoplasma capricolum*, Parainfluenza 3 virus and viruses of the Retroviridae family. Some of the bacteria listed here can also commonly be found in the upper respiratory tract.

For example, *P.haemolytica* is a common commensal of the ovine upper respiratory tract (121, 122) and is a leading cause of bronchopneumonia in sheep. A study looking at the microbiota of the nasal cavity and tonsils of domestic and Bighorn sheep identified several common commensal bacteria, including the potential pathogens *P.haemolytica*, *P.multocida*. *S.aureus* and *A.pyogenes* (123). It has been suggested that lung infections caused by *P.haemolytica* may be due to outgrowth from the upper respiratory tract, rather than by the inhalation of pathogenic strains, as *P.haemolytica* serotypes isolated from pneumonic sheep lungs usually correlate with those colonising the same sheep's nasal passages (124).

Lung infection by other microorganisms can change the likelihood of a sheep developing respiratory disease. For example, the presence of viruses in the lungs pre *P.haemolytica* infection has often been linked with poorer clinical outcomes (125-128) and the presence of *M.ovipneumoniae* has been shown to predispose Bighorn sheep to fatal *P.haemolytica* infections (129). On the other hand, co-infection by either *Bibersteinia trehalosi* or *P.multocida* may lead to better clinical outcomes as these bacteria are able to inhibit the growth of *P.haemolytica* in a proximity-dependant manner (130, 131).

Attempting to change the sheep respiratory microbiota through the use of probiotics or antibiotics to promote 'disease-preventing' microbial communities therefore seems like a reasonable goal. While I am not aware of any previous research that compares the composition of respiratory tract microbiota pre and post antibiotic treatment in sheep, this research has been done in cattle, which share many members of their respiratory microbiota with sheep. In both healthy calves and calves with respiratory disease, Allen et al. found that *P.multocida*, *P.haemolytica*, *Haemophilus somnus*, *Mycoplasma bovis* and *Mycoplasma bovirhinis* could be commonly isolated from the lower airways pre-antibiotic treatment (132). Subsequent antibiotic treatment (penicillin +/- trimethoprimsulfadoxine) did cause a decrease in *Pasteurella* spp. and *H.somnus* but for the *Pasteurella* spp. this effect was short lived as after the cessation of treatment rapid recolonisation of the respiratory tract was observed for these bacteria. This echoes the findings of Li et al. (74) in cystic fibrosis and may indicate that causing a long term change in the respiratory tract using antibiotics to promote lung health may not be possible. However, this meta-analysis by Li et al. analysed samples from 18 different studies which used varying methodologies, both in terms of lung microbiota sample collection techniques (bronchoalveolar lavage and sputum collection) and 16S analysis. Despite the cumulatively large sample size generated by combining data from these studies, caution should always be taken when

comparing 16S sequence data which has been produced using different methodologies as these can introduce differing biases into datasets (see Section 1.3).

Further research is required to characterise the sheep lung microbiota before therapeutic attempts to manipulate its composition can be attempted. 16S rRNA gene sequencing is a method by which we can research the bacterial communities of the lung without the biases associated with traditional culturing methods. This method and the biases associated with it are discussed further in the following chapter.

1.3 Sequencing low biomass bacterial communities

1.3.1 The 16S rRNA gene

A commonly used method for the identification of bacterial communities by non-culture based methods is through 16Ss rRNA gene sequencing. The 16S rRNA gene is ubiquitous amongst bacteria and contains nine variable regions flanked by conserved regions which can be used during primer design (**Fig. 1.3**). These variable regions can be sequenced in order to allow the taxonomic identification of the bacteria from which they originate. It is important to note that while using such a method removes the bias towards culturable microbes, it introduces new forms of bias which I will briefly describe below.

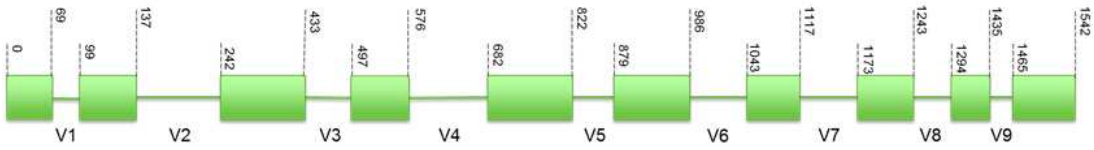


Figure 1.3: The 16S rRNA gene with the base pair locations of the variable regions indicated by their common identifier (V1-V9). Base pair numbering is based on the *Escherichia coli* system of nomenclature. Data adapted with permission from (133).

The copy number of the 16S rRNA gene per cell has been found to range from 1-15 copies per cell, depending upon bacterial taxonomy, although copy numbers of above 7 are only rarely identified (134). This means that, in a given sample, the relative abundances of 16S genes from multiple bacterial species are unlikely to be equivalent to the relative abundances of these bacteria in the original sample population. Also, as it is not possible to sequence the entire 16S rRNA gene using most sequencing technologies, it is necessary to select a specific region of the gene to amplify and sequence. This can be a source of bias, with several commonly used primers having been associated

with non-coverage of certain bacterial groups (135). It is unlikely that any variable region will be completely free of bias but some regions have been shown to provide generally good taxonomic coverage and depth, such as the V1-V2 and V2-V3 regions (133, 136, 137). Despite these sources of bias, as long as consistent protocols are used it is still possible to accurately and quantitatively measure changes in the relative abundances of different bacterial taxonomies between groups (138).

1.3.2 The impact of contamination

As well as the general difficulties arising from using 16S rRNA gene sequencing to study the microbiota, there are specific problems related to studying low biomass microbial communities such as those found in the lung. The most important of these is the increased impact of contamination, which if not appreciated and controlled for can lead to spurious results**.** Studies amplifying DNA from small quantities of bacteria risk amplifying DNA from contaminating bacteria along with genuine members of the bacterial community being studied. Contamination may originate from various sources including reagents, laboratory equipment or lab users (139).

The lower the level of bacterial template contained in the original sample, the greater the proportion of bacterial sequences originating from contamination will be after sequencing. This has been demonstrated in a study by Salter et al. which sequenced the 16S rRNA genes of serial dilutions of a pure *Salmonella bongori* culture (five dilutions of 10^8 to 10^3 cells) which were amplified by PCR for either 20 or 40 cycles prior to sequencing (140). For both 40 and 20 PCR cycles, when the template was reduced to 10^3 bacterial cells the majority of sequences were derived from contamination. Contaminants were found to vary between laboratories and DNA extraction kits (by manufacturer and lot number). Studies examining the lung microbiota usually find around 10^3-10^5 16S rRNA gene copy numbers per µl in lung samples (19, 20, 49), therefore it is possible that a large proportion of the sequences generated during these studies will originate from contamination (a separate study by Biesbroek et al. found that contamination began to have an effect on the apparent microbiota of a saliva sample when the sample contained < $1x10^{-6}$ ng/µl of template DNA (141)). Common contaminants identified by Salter et al. included soil and water bacteria, many of whom are nitrogen fixing; the authors suggest that this may be due to the use of nitrogen in ultrapure water storage tanks. Background contamination from extraction kit controls was found to be around 500 16S rRNA gene copies per µl while a separate study by Segal et al. found that in saline bronchoscope wash controls contamination reached the level of 41,195 16S gene copies per µl (49).

This level of contamination potentially explains some of the results from studies which have examined environments not normally considered to have a microbiota, such as the placenta, amniotic fluid, meconium, the surface of the eye and the brain (142-145). These studies frequently identify environmental nitrogen fixing bacteria such as *Methylobacterium*, *Nitrobacter* and *Beijerinckia* which may arise from contamination. A recent, well-controlled study examining placental samples

concluded that the placenta did not have a distinct microbiota as placental samples contained predominantly soil and plant associated bacteria which were also found in extraction kit reagent controls (146). It is important to keep in mind the effect of contamination while performing lung microbiota studies and to be cautious about any bacteria identified which are known to be nitrogen-fixing.

1.3.3 Controlling contamination and bias in low biomass samples

While it is often not possible to eradicate all possible sources of contamination there are ways of reducing it. Sources of contamination which are specific to lung microbiota studies are the upper respiratory tract (a problem when collecting sputum samples) and the bronchoscope channel, both of which can be avoided through the use of a sampling method which does not come into contact with these surfaces. BAL is a common method for sampling the lung which involves flushing phosphate buffered saline (PBS) or saline through the bronchoscope channel into the lung then collecting the fluid back through the channel. Unless saline flushes of the channel are included prior to each lung sampling, potentially leading to double the amount of samples needing to be sequenced, it is not possible to analyse the effect that the contaminated channel may have had on the lung samples (147). Protected specimen brushings, which do not come into contact with either the upper respiratory tract or the bronchoscope channel and thereby avoid contamination from these environments (29), can be used to avoid the need for these controls (see Section 2.5.3).

Contamination can also arise from the equipment and reagents used to collect samples, perform DNA extractions and amplify DNA by PCR. The quantity of contaminating bacterial DNA from these sources can be reduced by various techniques, including the use of standard sterile techniques and ultra violet (UV) treatment of reagents, workstations and equipment. However, UV treatment is unlikely to destroy all of the contaminating DNA (148) and it is not possible to UV certain reagents such as deoxynucleotide triphosphates or taq polymerase (149). Other methods have been tried, including deoxyribonuclease treatment, restriction endonuclease digestion and DNA intercalation (150-152) but if the compounds used in these methods are not completely inactivated or removed they may inhibit PCR reactions. Other researchers sharing the same lab space have also been suggested as a risk factor for contamination (153), but obviously this is often unavoidable.

In order to further reduce the quantity of contaminating DNA on surfaces and equipment, it is possible to combine other techniques with chemical treatment. Bleach and copper-bis-(phenanthroline)-sulfate/H_20_2 solution have been found to be the most effective chemical treatments but as these compounds are corrosive and therefore cannot be used on certain materials, commercial products such as DNA away by Molecular Bioproducts have been suggested as an adequate alternative (149).

In conclusion, there is currently no generally accepted method which will completely eradicate all contaminating DNA while not affecting PCR efficiency or damaging reagents/equipment and it is

therefore highly important to process reagent only controls alongside samples when carrying out studies on the lung microbiota, allowing the researcher to identify which bacterial sequences in samples may be due to contamination. As of yet there is also no generally accepted method for removing contaminating sequences from datasets although methods for doing so have been suggested including identifying likely contaminants using the neutral model (39) or by combining sequencing information from negative controls with quantitative PCR (qPCR) data (154). Other than contamination, the largest source of bias related to sequencing low biomass bacterial communities is the large amount of PCR cycles which are needed to amplify the small quantity of DNA in the samples. This can affect the apparent diversity of the bacterial communities sampled and can lead to increased taxonomic bias (155, 156). Low PCR template concentrations (0.1 ng per reaction) have also been shown to give lower amplification reproducibility in comparison to higher template concentrations (5 to 10 ng per reaction) (157). While it is likely to be impossible to eradicate PCR bias completely, the inclusion of a mock community control made up of known quantities of a mixture of bacterial species can help identify any biases which may be present (158, 159).

1.4 Objectives

The overall aim of my thesis is to characterise the lung microbiota of sheep in order to assess its potential usefulness as a model of the human lung microbiota or to study sheep respiratory disease. The following outline provides a brief description of the objectives of each of my results chapters.

- Chapter 3: Variability of the Lung Microbiota – To analyse the amount of inter- and intra-individual variability present in the sheep lung microbiota (n=6 animals). Also, to offer a better understanding of the types of bacteria which can be found in the sheep lung.

- Chapter 4: Comparing Microbiotas in the Upper Aerodigestive and Lower Respiratory Tracts of Lambs – To characterise the lung microbiota composition in a larger number of animals (n=40). Also, to discover whether, as in humans, the sheep lung microbiota is similar to that of the upper respiratory tract.

- Chapter 5: Microbiota in Exhaled Breath Condensate and the Lung – To assess whether it is possible to replace invasive lung sampling methods with the less invasive sampling method of exhaled breath condensate (EBC) collection. Also, to discover whether it is possible to manipulate the lung microbiota using nebulised antibiotic treatment.

- Chapter 6: Identifying Core Members of the Sheep Lung Microbiota – To analyse the inter-study variability of my results and to attempt to identify bacteria commonly isolated from the sheep lung in all of my studies to identify a 'core' sheep lung microbiota.

The following methodology chapter covers the materials and methods I used during my results chapters and the reasoning behind their use.

Chapter 2: Methodology

2.1 Ethical statement

Experimental protocols were subject to the Animals (Scientific Procedures) act 1986 and all studies involving animals were approved by the Roslin Institute Animal Welfare and Ethics Committee. A statistician (Dr. Helen Brown) was consulted to advise on appropriate sample sizes.

2.2 Sheep details

Table 2.1 contains the available information about the sheep used in each results chapter. Due to the outbred nature of the sheep used in my thesis, genetic variation is likely to be high and it is possible that this variation will impact the types of bacteria which can be found as members of the lung microbiota.

Table 2.1: Sheep used in this thesis

Chapter	Gender	Age	Breed	Weight (mean ± SD at first sampling)
3 (contains two studies, one using n=6 sheep and one using n=1)	N=6: 5 females and 1 castrated male N=1 Female	N=6: 20 months old N=1: 36 months old	Suffolk-cross	N=6: 60.3 ± 7.3 kg N=1: 60 kg
4	20 females and 20 males	48.8 days (mean) ± 0.8 (SD)	Scottish Mule X Suffolk	20.6 ± 2.6 kg
5	6 castrated males	14 months	Suffolk-cross	49.2 ± 3.4 kg

2.3 Removal of contaminating DNA

Due to the low quantities of DNA present in the healthy lung, it is important to reduce the levels of potentially contaminating bacterial DNA from reagents and equipment used to process samples. During DNA extraction and PCR setup, all experiments were carried out in either a safety cabinet or a PCR cabinet to prevent contamination from aerosolised bacteria. Prior to use, cabinets were either treated with 30 mins of UV radiation or with Ambion DNAZap PCR DNA Degradation Solutions (Ambion inc., USA). All pipettes, tubes and reagents used during DNA extraction, PCR and qPCR, were UV treated in a Hoefer UVC 500 Crosslinker (Hoefer inc., USA) except for the DNA template, deoxynucleotide triphosphates and the Q5 Hot Start High-Fidelity 2X Master Mix (New England Biolabs, UK). UV treatment has previously been shown to be effective at decreasing contaminating DNA of the length which is generated by the first step in my PCR protocol (160).

2.4 Anaesthesia

Before anaesthesia, food was withheld from animals for 24 hours. Anaesthesia was induced by injecting 6-8 mg/kg of propofol (Fresenius propofol, 1%, Fresenius Kabi Ltd, UK) and maintained by positive pressure ventilation using a 50:50 mix of nitrous oxide and oxygen with 1-3% isoflurane. Tidal volume was adjusted to 10 ml/kg body-weight and end-tidal CO_2 at 4.5-5.5% by adjusting the respiration rate.

2.5 Sample collection

2.5.1 Bronchoscope wash controls

To act as a control for some bronchoscopic procedures, 7.5 ml sterile PBS was passed through the bronchoscope channel prior to sample collection (see Chapter 3). The wash was collected in a sterile tube and stored on ice for up to three hours.

2.5.2 Exhaled breath condensate (EBC) collection

For Chapter 5, EBC was collected from both conscious and anaesthetised animals using an RTubeVENT breath condensate collection device (Respiratory Research, USA). RTubeVENTs were treated with UV before use as they are not provided sterile. EBC was collected for 10 mins from conscious sheep held in a holding crate. A face mask was attached to the sheep and an RTubeVENT within an insulated cooling sleeve (cooled to ≤ -20°C) was attached to the face mask as shown in **Chapter 5 (Fig. 6)**. EBC was collected from anaesthetised sheep for 10 mins, as also shown in **Chapter 5 (Fig. 6)**, to bypass and reduce contamination from the upper respiratory tract.

After EBC collection, the RTubeVENTs were capped and transferred to a microbiological safety cabinet where an aluminium plunger was used to pool the EBC. This is performed in the following manner: the RTubeVENT is pushed down over the plunger which forces a piston at the bottom of the tube up through the tube, gathering the condensate from the tube along its rim and pushing it towards the top of the tube. A second piston at the top of the tube is removed which allows the condensate collected around the edge of the first piston to be accessible. The condensate is collected using a pipette and transferred to an Eppendorf tube.

2.5.3 Protected specimen brushings (PSBs)

For Chapters 3 and 5, brushings were taken using a protected specimen brush (ConMed Disposable Microbiology Brush, Conmed, USA) to prevent contamination of lung samples by the upper airways and to allow greater spatial precision when taking samples from the lung epithelium (**Fig. 2.1**).

The bronchoscope was positioned at the sampling site then the protected specimen brush catheter was advanced 1-3 cm from the bronchoscope and the protective carbon wax plug was ejected into the lung where it will have safely dissolved. The brush was then advanced from the catheter to the epithelium; the sampling area was brushed thoroughly before the brush was retracted into the catheter sleeve and the catheter was removed from the bronchoscope. Brushes were cut using sterile scissors into 1 ml PBS so that the bacteria on the brushes would diffuse into the PBS. After centrifugation for 15 mins at 13,000g (Biofuge Fresco**,** Heraeus, Germany**)**, a portion of the supernatant was removed, leaving 500 µl fluid.

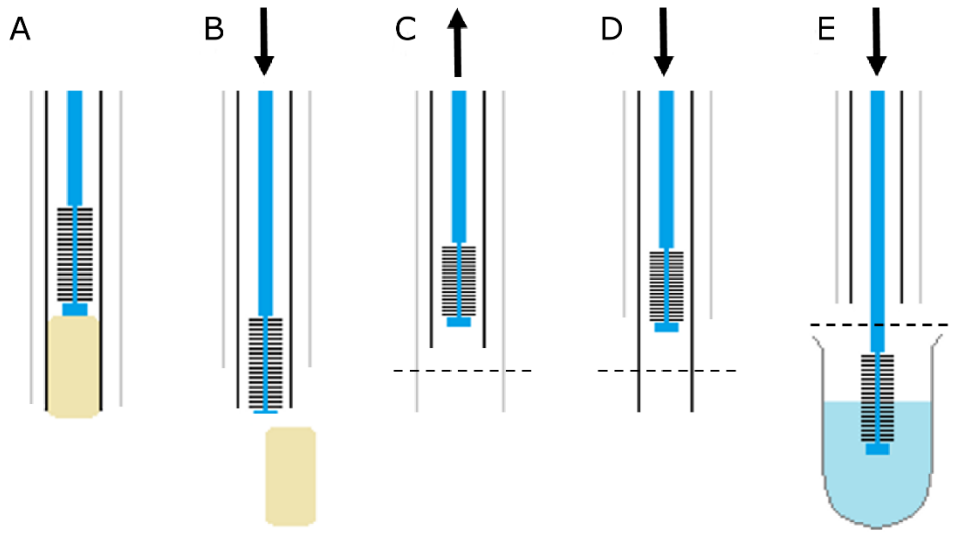


Figure 2.1: Protected specimen brushings of the lung epithelium. A) The protected specimen brush sits within a catheter, protected from contamination by the presence of a carbon wax plug. B) Once the catheter is positioned near the sampling site, the inner catheter is advanced forward. This ejects the wax plug into the lung where it safely dissolves. The brush is then advanced and a sample is taken from the lung epithelium. C) The brush and inner catheter are withdrawn inside the outer catheter and removed from the bronchoscope. The sample is then transferred to a safety cabinet where the end of the outer catheter is removed using sterile scissors (dashed line). D) The inner catheter is advanced fully and the brush is retracted 5mm within the inner catheter. The inner catheter is cut 5mm from the tip (dashed line). E) The brush is advanced from the catheter and is cut using sterile scissors into 1 ml PBS.

2.5.4 Bronchoalveolar lavage (BAL)

The bronchoscope was advanced and wedged in the segment to be sampled then 2 x 20 ml PBS was used to collect BAL fluid which was transferred into sterile tubes (through sterile gauze in order to remove mucus which would have interfered with the quantification of colistimethate sodium by high performance liquid chromatography) and immediately stored on ice (see Chapter 5).

2.5.5 Throat swabs

For Chapter 4, cotton tipped swabs contained within a protective plastic sheath (Swab Plain Wood Cotton Tip Sterile, Copan, Italy) were used to prevent contamination. The swab was retracted within the sheath and the end of the sheath (closest to the swab tip) was removed with sterile scissors. The end of the sheath was placed at the back of the mouth and the swab was advanced to sample the throat before being retracted into the sheath prior to removal from the mouth (**Fig. 2.2**). The swab was transferred into a fresh sheath before being stored on ice. The swab tip was cut into 500 µl of sterile PBS for storage.

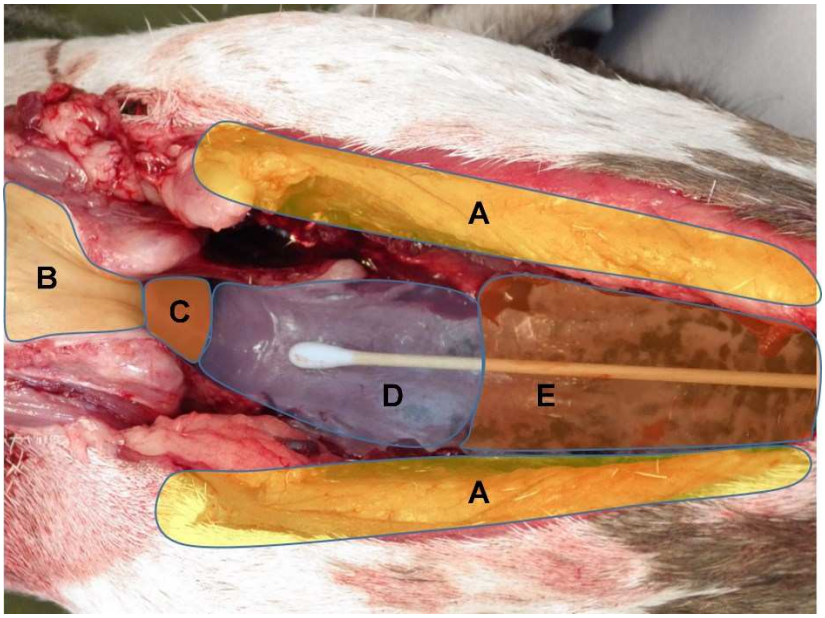


Figure 2.2: Throat swab sampling site. A sheep head dissected to show the position in the throat from which throat swabs were taken. The swab did not pass the soft palette and can therefore be said to have sampled the oropharynx and not the nasopharynx. A) Ventral aspect of mandibles. B) Tongue (reflected back). C) Ventral aspect of epiglottis. D) Soft palette. E) Hard palette.

2.5.6 Collection of lamb lung fluid by syringe

For Chapter 4, the ventral aspect of the neck of the animals was shaved before the ventral trachea was exposed using a scalpel. A sampling site on the trachea was identified and an area proximal to this site was tightly closed using string and a clamp. The sampling site was seared using a heated spatula before 50 ml PBS was injected into the tracheal lumen. The head and neck were raised to encourage flowing of the PBS caudally into the lower airways then a clamp was placed below the site of injection to prevent leakage from the injection site. A site on the thorax caudal to the clamp was seared as above and the sheep was tipped so that the fluid flowed into the tracheal lumen to enable maximum fluid collection. Fluid was collected through the seared section using a syringe and transferred to sterile tubes.

2.6 Sample storage

Samples were stored on wet or dry ice within 30 mins of collection and were then stored at -80°C or underwent DNA extraction within 4 hours. Storage for this length of time has previously been shown to have minimal impacts on sputum bacterial communities (161). Extracted DNA was stored at -80°C. Several studies have found that storage of samples at -80°C has little impact on the apparent composition of bacterial communities (162, 163) whereas others have shown a significant impact on the ratios of different bacterial groups (164, 165).

In Chapter 3, samples from the baseline and 1 month time-points were stored at -80°C then underwent DNA extraction within 1 month. Samples from the 3 month time-point underwent DNA extraction within four hours of sample collection. Samples from the 36 month old sheep also underwent DNA extraction within four hours of sample collection. In Chapter 4, samples were stored at -80°C for nine months prior to DNA extraction. In Chapter 5, samples were stored at -80°C and underwent DNA extraction within one month.

2.7 Colistimethate sodium (CMS) treatment

In Chapter 5, treatment of sheep with CMS is used to assess whether there is a quick, simple and non-invasive way to influence the sheep lung microbiota composition. CMS is one of only a small number of antibiotics which have been specifically formulated and manufactured for aerosolised use, which is a more efficient way of delivering antibiotics to the lung (116). It is commonly used in its aerosol form to treat exacerbations in cystic fibrosis caused by Gram negative pathogens. Colistin is a polymyxin type antibiotic which only has activity against Gram negative bacteria; CMS is a less toxic prodrug of colistin which is hydrolysed to colistin in aqueous solution.

Information on CMS dosing is often confusing due to the use of different units by manufacturers (166, 167) and the need for better information on the pharmacokinetic-pharmacodynamic relationship of colistin in different patient populations (168). The usual dose for humans over 2 years of age is 1-2 million units two to three times per day (maximum 6 million units per day) (169); unlike colistin, CMS has been shown to have minimal renal toxicity in humans, even at very high doses (9,000,000 units twice daily (170)).

As we intended this to be a simple method for changing the lung microbiota composition, we designed our dosage regimes to require no 'out of hours' animal work and for one course to not last longer than a week. Two million international units (IU) of CMS (Colomycin Injection, Forest Laboratories UK ltd., UK) was dissolved in 4 ml distilled water and added to the nebulisation chamber of an eFlow Rapid Nebuliser System (PARI, UK). Sheep were held in a holding crate and were fitted with a face mask which was connected to the nebuliser via plastic tubing fitted with valves to ensure that breathing only occurred unidirectionally. Doses were delivered until both the full dose was delivered and the machine automatically switched off or until 15 mins had elapsed, so that the sheep were not overly stressed due to longer periods of restraint.

Two dosage regimes were used:

- Dosage regime 1: A single dose of 2,000,000 IU CMS per day was administered. This was repeated every day for seven days after which blood and BAL fluid samples were collected 24 and 48 hours after the final dose. This dosage regime was used to establish that we were able to successfully administer a bactericidal dose of colistin to the lung using nebulisation (see Section 5.1).

- Dosage regime 2: Two doses of 2,000,000 IU CMS were administered with an interval of 6 hours. Blood and BAL fluid samples were collected 24 hours after the final dose. This dosage regime was used to test whether it was possible to cause a change in the lung microbiota using a single dose of nebulised antibiotic (see Section 5.2).

Due to the lack of controlled clinical trial information and the fact that different brands of CMS have been shown to lead to different exposures to the active form of colistin in rats (171), we decided to measure the quantity of colistin in the treated sheep after treatment. Serum was produced by centrifuging blood for 15 mins at 2500g (Megafuge 1.0, Heraeus, Germany). As the concentration of urea in epithelial lining fluid and serum should be equal (172) it is possible to find the dilution factor of the epithelial lining fluid to PBS in BAL fluid samples by calculating the difference between serum and BAL fluid urea concentrations. The urea concentrations in serum and BAL fluids were quantified by the Easter Bush Clinical Pathology Unit. Colistin was quantified in BAL fluid using High Performance Liquid Chromatography (HPLC) (performed by Andy Gill), using previously described methods (173, 174) (methods described in more detail in Chapter 5). The calculated colistin

concentration was then multiplied by the BAL fluid dilution factor in order to calculate the concentration of colistin in the epithelial lining fluid.

2.8 Differential cell counting

It is possible that the presence of infection in the lungs may lead to changes in the lung microbiota. To check whether infection is present, BAL fluid samples can be taken and the immune cells present counted. A high eosinophil count may indicate helminth infection while a high neutrophil count can also be a sign of infection.

BAL fluid was centrifuged at 1400g for 5 mins (Megafuge 1.0, Heraeus, Germany) and the supernatant was poured off leaving a small amount of liquid and a cell pellet. The pellet was resuspended in the remaining liquid and up to 2 ml PBS, depending on the size of the pellet. An improved Neubauer cell counting chamber and a light microscope were used to calculate the total white blood cell count and the live/dead white blood cell ratio. 10 µl nigrosin (Nigrosin water soluble, Sigma, UK) was mixed with 10 µl of the cell suspension and added to the haemocytometer. Nigrosin stains dead cells dark grey/black and can therefore be used to differentiate live from dead cells. 10 µl white blood cell counting fluid (made in house by Steven Wright using glacial acetic acid and crystal violet) was mixed with 10 µl of the cell suspension and added to a separate section of the haemocytometer. This counting fluid lyses red blood cells but not white blood cells.

Before loading the samples into the ctyospin (Ctyospin 2, Shandon, UK) the cell solution needed to be diluted to 5 x 10^4 cells/100 µl. The total quantity of fluid needed for the cytospin was calculated using the following equation:

$$\frac{200}{(Total\ white\ blood\ cell\ count\ \div\ 5 \times 10^4)} = \text{µl cell susupension}$$

To this was then added 200 µl of PBS. Two cytospin slides per sample were prepared by adding 100 µl of this solution to each slide, with the cytospin machine set for 5 mins at 600 rpm. Slides were allowed to air dry. They were then stained in 2 ml Leishman's stain (Leishman Stain in 100% Methanol, Fisher Chemical, UK) for 2 mins then in this plus 2 ml water for a further 8 mins. Slides were thoroughly washed in water, allowed to air dry then mounted in DPX mountant (DPX Mounting Medium, Agar Scientific, UK) before differential cell counting using a light microscope. Cells were classified as alveolar macrophages, neutrophils, mast cells, eosinophils, monocytes, lymphocytes and epithelial cells according to standard morphological criteria (175). The normal cellular composition of healthy sheep BAL fluid is <7.5 x 10^6 cells/ml, >77.7% alveolar macrophages, <9.5% neutrophils, <7.5% eosinophils, <0.6% mast cells, <11.2% lymphocytes and <3.4% of other cell types (Personal communication, Steven Wright and David Collie). Differential cell counting was performed for all sheep in Chapter 5. Differential cell counts were found to be within normal ranges except for sheep

ED952 which had an elevated eosinophil count (19.2%) and sheep ED956 which had a high lymphocyte count (19.6%).

2.9 DNA extraction

Different DNA extraction kits and methodologies can affect not only the quantity of DNA but also the apparent composition of the bacterial communities extracted from samples, making it difficult to compare results obtained using different methods (176, 177). I have therefore used the PowerSoil DNA Isolation Kit (Mo Bio, USA) and the same extraction protocol for all experiments. A bead beating step is included in this protocol as bead beating has been shown to produce a more representative bacterial community structure (178, 179).

The PowerSoil DNA Isolation Kit was used according to the manufacturer's instructions, except where adjustments were made to facilitate the processing of different sample types; an added heating step prior to bead beating to increase cell lysis and the use of less DNA eluent to increase the final DNA concentration. Extraction kit reagent controls were produced using the same protocol to identify any potential bacterial DNA originating from the kit. All centrifugation steps were carried out in a Biofuge Fresco **(**Heraeus, Germany**).**

For protected specimen brushes, lamb lung fluids, and PBS controls: Samples were briefly vortexed and centrifuged at 13,000g for 5 mins. The supernatant was removed leaving any cell pellets, brushes and ~100 µl of liquid. 60 µl of PowerSoil Solution C1, a cell lysis agent, was added to each sample and was used to resuspend pelleted cells. The suspension and brushes (where present) were transferred to bead beating tubes along with 750 µl of PowerSoil Bead Solution. Swabs were placed directly in bead beating tubes containing 750 µl PowerSoil Bead Beating Solution and 60 µl PowerSoil Solution C1.

Samples were then heated at 65°C for 15 mins before being placed in a FastPrep FP120 Cell Disrupter for 45 s at 5.0 m/s. This heating step was included to increase cell lysis. Tubes were centrifuged at 10,000g for 30 s and 250 µl of Powersoil Solution C2 was mixed with the supernatant in separate 2 ml tubes. Samples were incubated at 4°C for 5 mins then centrifuged at 10,000g for 1 min. 600 µl of the supernatant was transferred to new 2 ml tubes and mixed with 200 µl Powersoil Solution C3. Solutions C2 and C3 both act to precipitate any non-DNA organic and inorganic material from the samples. Samples were again incubated at 4°C for 5 mins then centrifuged at 10,000g for 1 min. The supernatant was transferred to new 2 ml tubes and mixed with 1.2 ml of Powersoil Solution C4 (a concentrated salt solution which will allow DNA binding to silica). This was then centrifuged through a silica filter at 10,000g for 1 min followed by 500 µl of Powersoil Solution C5 (an ethanol based wash solution) at 10,000g for 30 s. The filter was then centrifuged at 10,000g for 1 min to remove any residual Solution C5. 50 µl PowerSoil Solution C6 (eluent) was added to the silica filter and incubated

at room temperature for 5 mins before centrifugation at 10,000g for 30 s. The resulting fluid was collected in separate 2 ml tubes and the DNA concentrations were measured using a Nanodrop 2000 (Thermo Fisher Scientific, USA).

2.10 16S rRNA gene amplification

The primers I used for 16S rRNA gene amplification were originally chosen as they would produce a ~450 bp product covering the V2-V3 regions of the gene which it was thought would give greater taxonomic depth than using a shorter amplicon as using sequences of >250 bp was recommended for the best taxonomic coverage (137, 180). At the time they were designed it was not appreciated that as the Illumina MiSeq could only produce 250 bp paired end reads the overlap between the reverse and forward reads for a 450 bp sequence would be very small which would lead to bioinformatic programs being less able to detect sequencing errors in the non-overlapping regions. This increases bioinformatic processing time and also increases the sequencing error rate. As described in my introduction (Section 1.3.1), primer choice can greatly influence the apparent composition of the microbiota. The primers were designed and used by my group to produce sequence data on the sheep lung microbiota prior to my PhD (115) and in order to be able to compare my data with these findings it was decided that I would continue to use them, as long as the error rate was not found to be high. A two-step nested PCR protocol was used to reduce potential bias caused by sequence barcodes, which increases in relation to the number of PCR cycles (181). However, it must be noted that a high number of PCR cycles was still used and this has been related to increased PCR bias (155, 156).

For both PCR steps, 25 µl of Q5 High-Fidelity 2X Master Mix (New England Biolabs) and 2.5 µl of each primer (10 µM) were used. A positive and negative control was included during each PCR run. The positive control consisted of DNA extracted from *P.aeruginosa* Strain PA0579 and the negative control consisted of nuclease free water (Qiagen, USA). Mock communities were included to control for PCR or sequencing bias. In Chapters 3 and 4 the Human Microbiome Project mock community HM-782D (Genomic DNA from Microbial Mock Community B (Even, Low Concentration), v5.1L, for 16S rRNA Gene Sequencing) was used while in Chapter 5 the Human Microbiome Project mock community HM-783D (Genomic DNA from Microbial Mock Community B (Staggered, Low Concentration), v5.2L, for 16S rRNA Gene Sequencing) was used. HM-782D contained genomic DNA from 20 bacterial species with 100,000 ribosomal RNA operon copies per organism per µl. The twenty bacterial species included were *Acinetobacter baumannii*, *Actinomyces odontolyticus*, *Bacillus cereus*, *Bacteroides vulgatus*, *Clostridium beijerinckii*, *Deinococcus radiodurans*, *Enterococcus faecalis*, *Escherichia coli*, *Helicobacter pylori,*, *Lactobacillus gasseri*, *Listeria monocytogene*, *Neisseria meningitides, Propionibacterium acnes*, *Pseudomonas aeruginosa, Rhodobacter sphaeroides*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Streptococcus agalactiae*, *Streptococcus mutans* and *Streptococcus pneumoniae*. HM-783D contained the same species but there

were 1,000 to 1,000,000 ribosomal RNA operon copies per organism per µl. Both mock communities were supplied by BEI Resources.

The first PCR step amplified the V1-V4 region of the 16S rRNA gene using the primers 28F (5'-GAGTTTGATCNTGGCTCAG-3') and 805R (5'-GACTACCAGGGTATCTAATC-3'). The PCR conditions were: 94°C for 2 mins followed by 20 cycles of 94°C for 1 min, 55°C for 45 s and 72°C for 1.5 mins followed by 72°C for 20 mins. Amplicons were purified using the AMPure XP PCR Purification system (Beckman Coulter, UK) (described in Section 2.11). Amplicons were eluted into 20 µl nuclease free water.

The second PCR step amplified the V2-V3 region of the 16S rRNA gene using the barcoded primers 104F (5'-GGCGVACGGGTGAGTAA-3') and 519R (5'-GTNTTACNGCGGCKGCTG-3'). These were Integrated DNA Technologies primers with Trugrade processing and Truseq (i5 and i7) and Nextera (i5) index barcodes. The PCR conditions were: 98°C for 30 s followed by 20 cycles of 98°C for 10 s, 67°C for 30 s and 72°C for 10 s followed by 72°C for 2 mins. The resulting amplicons were purified using the AMPure XP system and amplicons were eluted into either 30 µl nuclease free water or 0.1% Tween20 in Qiagen Elution Buffer. Amplicons were viewed on a 1.2% agarose gel (made in house using TAE buffer and agarose (Sigma, UK)) using SYBR Safe DNA Gel Stain (Thermo Fisher Scientific, UK) and a UV illuminator. Labels were affixed to gel images using ImageJ (http://imagej.net/).

2.11 Ampure purification

To purify amplicons from PCR mixtures I used the AMPure XP PCR Purification system. This method allows the purification of DNA of above a user specified length by the reversible binding of the DNA to beads coated in carboxyl molecules. This method of purification is recommended by Edinburgh Genomics (https://genomics.ed.ac.uk/resources/sample-requirements).

The PCR mixture was mixed with the Ampure reagent by pipetting 10 times. At first I used a 1.8:1 ratio of beads to sample fluid, as recommended by the manufacturer, but at a later stage in my PhD (for results found in Chapter 5) this was changed to a 1:1 ratio which was found to be equally efficient at purifying the amplicons. This mixture was then incubated for 5 mins at room temperature then placed on a magnet for 2 mins to pellet the beads. The liquid was aspirated and the beads were washed three times with 200 µl 70% ethanol (30 s incubation followed by aspiration for every wash). After the final aspiration, beads were allowed to air dry for 5 mins. The samples were then removed from the magnet and 30 µl of eluent was mixed with the beads, followed by 5 mins of incubation at room temperature. The samples were then placed onto the magnet for 1 min and the eluent containing the DNA was removed.

2.12 Sequencing strategy

2.12.1 Library preparation

It is necessary to ensure that equimolar quantities of each sample are added to the sequencing library. Samples were pooled into sequencing libraries by either Edinburgh Genomics or by myself using the protocol described below.

The concentration of DNA in PCR amplicons was calculated using a Qubit 3.0 Fluorometer (Life Technologies) and the Qubit dsDNA HS Assay Kit (Life Technologies) according to the manufacturer's instructions. A working solution was created by diluting the Qubit dsDNA HS Reagent 1:200 in Qubit dsDNA HS Buffer. 190 µl of the working solution was mixed with 10 µl of the two Qubit Standards and 198 µl of the working solution was mixed with 2 µl of each sample. All samples and standards were allowed to incubate at room temperature for >2 mins to 1 hour. On the Qubit 3.0 fluorometer, the dsDNA High Sensitivity assay was selected and the standards were read prior to every batch of sample readings. The concentration of each sample was outputted as ng/µl and this was used to calculate the nM concentration.

Equimolar amounts of samples were pooled, except where the amount of sample to be added to the library would have been greater than 10 µl in which case 10 µl of sample was added to the pool. All DNA pools contained >5 nM of DNA. Pooled libraries were quality tested using High Sensitivity D1K ScreenTape (Agilent Technolgies) by Edinburgh Genomics (**Fig. 2.3**).

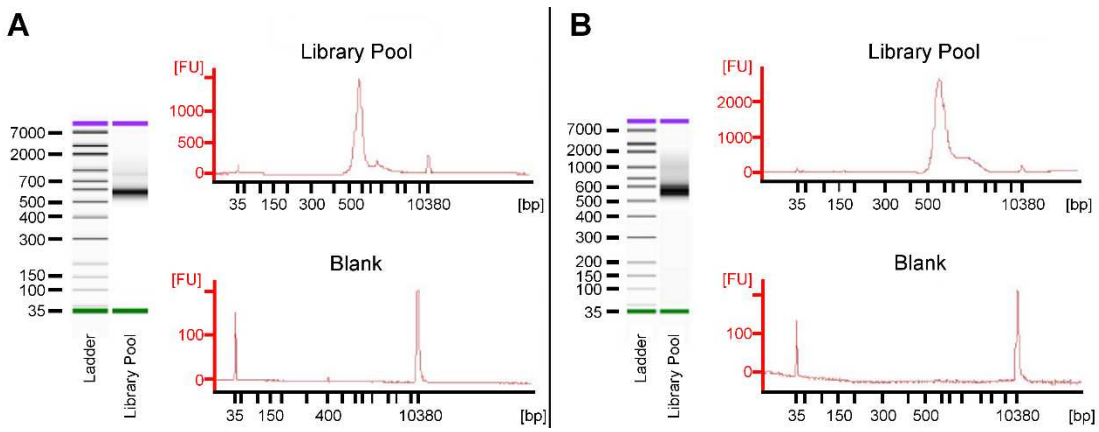


Figure 2.3: Comparison of pooled sequencing libraries. DNA from sequencing libraries was visualised by Edinburgh Genomics using an Agilent Bioanalyser 2100 DNA 1000, High Sensitivity chip. In A) samples were pooled into a library by Edinburgh Genomics and in B) samples were pooled by me. Both libraries show a clear peak at around 550bp which was the expected fragment size. The pool in A) contained 5.6 ng/µl DNA and B) contained 3.32 ng/ µl.

2.12.2 Sequencing

Various sequencing platforms are available which can be used for 16S rRNA gene sequencing including the MiSeq/HiSeq (Illumina), Pacific Biosciences, Ion Torrent (Thermo Fisher Scientific), and 454 (Roche) platforms. Of these, the Illumina MiSeq has been shown to produce more accurate data (138, 182). The MiSeq was therefore selected as the optimal sequencer. However, due to a problem with a version of the Illumina chemistry which resulted in low quality reads for 16S rRNA gene sequences, some of my samples were sequenced using the Illumina HiSeq (some samples from Chapters 3 and 4). The HiSeq uses similar chemistry to the MiSeq but produces a higher number of reads and a lower error rate (183). Libraries were sequenced by Edinburgh Genomics using either the Illumina MiSeq or HiSeq2500 (Rapid mode) producing paired 250-nucleotide reads.

2.13 Quality control and processing of sequences

2.13.1 Primer removal

Primers were removed from fastq files using cutadapt (184) which has been shown to provide a good balance between sensitivity (ratio of properly trimmed reads to improperly trimmed reads) and specificity (ratio of correctly untrimmed reads to reads which were trimmed but should not have been) (185). 5' adapters were removed along with any sequences proceeding them. A maximum sequence error rate of 10% was allowed for primers which included errors due to mismatches, deletions and insertions. All sequences with higher error rates were discarded.

2.13.2 Quality control and OTU clustering

The mothur program (186) was used for quality control of sequences, alignment of sequences to a reference database, taxonomic assignment and OTU clustering, following a modified version of the MiSeq pipeline designed by the mothur creators (187). The mothur program has been shown to demonstrate similar or better results to other commonly used 16S rRNA gene bioinformatic pipelines such as QIIME and MG-RAST (188, 189).

Sequences were constructed from aligned forward and reverse reads. Where there was disagreement between the overlapping portions of the forward and reverse reads, if one base had a quality score 6 points greater than the corresponding base on the opposite read it was used in the final sequence. Where there was not a 6 point difference between the base quality scores the base was assigned an N. If there was a gap at the same position as a base then the base was only considered to be genuine if it had a quality score of ≥25 points. Any reads which did not have matching forward or reverse reads were discarded.

Sequences with ambiguous base calls, which were less than 369 bases in length or contained homopolymers of greater than 9 base pairs were discarded. Sequences were aligned against the SILVA database (190) as this has demonstrated better alignment quality than Greengenes, MUSCLE and RDP alignments and thereby creates less artificial sequence variation (180). Sequences which did not align with the V2-V3 region of the 16S rRNA gene were discarded. Chimeras were identified and removed using Uchime (191) within mothur.

Sequences were classified using mothur's Bayesian classifier and the Greengenes database (192) trimmed to the V2-V3 region of the 16S rRNA gene to improve classification depth (193). The Greengenes database was selected for sequence classification as it has been demonstrated to be able to identify a greater proportion of sequences than SILVA or RDP training sets (137, 193). However, the Greengenes database has not been updated since 2013 and this should be taken into account when interpreting my results (194). Sequences which did not originate from bacteria were discarded. OTUs were clustered by phylotype using a database dependent approach and were then sub-sampled prior to statistical analysis. Samples were subsampled to the lowest number of reads found in any of the samples within the study, prior to statistical analysis. The sequencing error rate was calculated by comparing a sequenced mock community control to a reference file containing the correct mock community sequences.

2.14 16S rRNA qPCR

Triplicate reactions of 20 µl were carried out in a LightCycler 480 (Roche), targeting the V3 region of the 16S rRNA gene. This protocol was designed by Ms Tina Baker. Primers were included at a final concentration of 0.4 µM. The reaction mixture consisted of 1 µl of extracted DNA solution, 0.2 µl of primer UniF340 (5'–ACTCCTACGGGAGGCAGCAGT–3'), 0.2 µl of primer UniR514 (5'–ATTACCGCGGCTGCTGGC–3'), 10 µl LightCycler® 480 SYBR Green I Master mix (Roche) and 8.6 µl water. After the addition of the reaction mixtures, qPCR plates were sealed and centrifuged at 15,000g for 2 mins (Megafuge 1.0, Heraeus, Germany). The PCR parameters used are described in **Table 2.2**. While a melting curve program was ran, this is of limited use due to the heterogeneous nature of the gene amplicons; abnormally shaped melting curves with multiple peaks are often observed due to the differing GC contents of different bacterial species' DNA.

Table 2.2: qPCR protocol for the V3 region of the 16S rRNA gene

Program	Cycles	Target (°C)	Hold (mins:s)	Ramp rate (°C/s)
Pre-incubation	1	50	02:00	4.80
		95	10:00	4.80
Amplification	45	95	00:30	4.80
		63	00:30	2.50
Melting curve	1	95	00:05	4.80
		65	01:00	2.50
		97	Continuous (Acquisitions per 1°C = 5)	0.11

A standard curve was constructed from amplicons generated by DNA extracted from a *P.aeruginosa* culture which had undergone the first round of PCR as described in Section 2.10, except that 30 cycles were performed rather than 20. These amplicons were purified using the Ampure system and their DNA concentration was measured using a Qubit 3.0 Fluorometer. A standard curve was constructed as shown in **Fig. 2.4**.

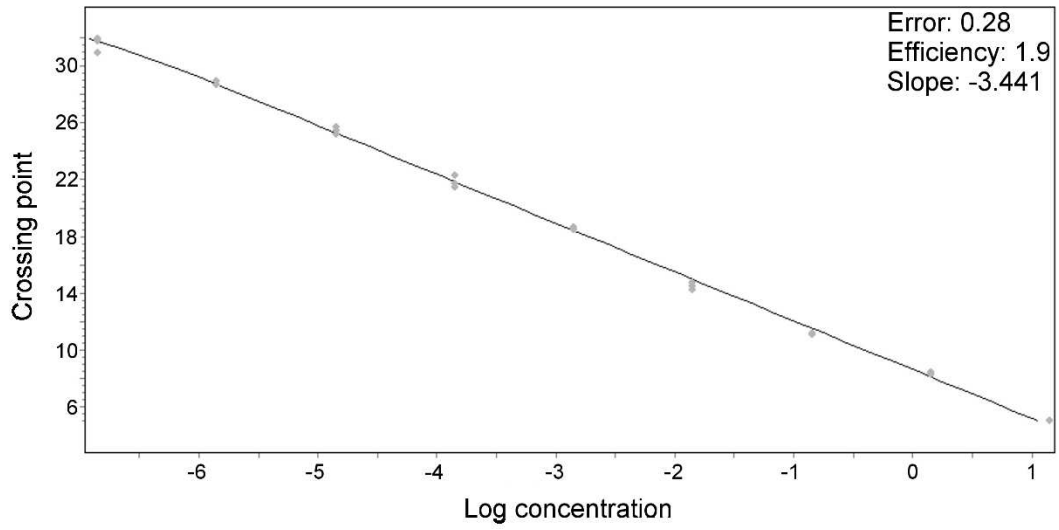


Figure 2.4: Example of a standard curve for 16S rRNA gene qPCR. Constructed using 777 bp 16S rRNA gene amplicons generated by PCR from DNA extracted from a *P.aeruginosa* culture.

2.15 Statistical and graphical analysis

Statistical tests were carried out within either mothur, R (R version 3.2.5), Minitab (Minitab 16 Statistical Software) or SPSS (IBM, SPSS Statistics 21). Graphs were constructed in either SPSS, Excel (Microsoft, Excel 2013), Minitab or R. GIMP 2.8 (www.gimp.org) was used to improve the image quality of graphs.

2.15.1 OTUs: Alpha-diversity

All tests were carried out within mothur.

Good's coverage

The Good's coverage indicator indicates what proportion of the total species present in a community is represented in a sample (195). A value of 1.0 would indicate that all of the species in the community were represented. I chose this as a means of estimating sequencing depth adequacy rather than rarefaction curves as it is a more easily interpretable measure of sequencing depth adequacy eg, if the Good's coverage estimator value is 0.99 this would mean that 99% of the bacteria in the original sample had been identified via sequencing.

Inverse Simpson diversity estimate

The Simpson diversity estimate estimates the likelihood that two sequences selected randomly from the same sample will belong to the same OTU. The inverse of this value can be used so that an increase in its value reflects an increase in diversity (196). I decided to use this measure of diversity as it is more weighted towards dominant OTUs. This means that any low abundance OTUs caused by sequencing errors which have not been removed during quality control will not greatly affect the observed diversity.

Chao1 richness estimator

The inverse Simpson diversity estimator is largely unaffected by low abundance OTUs and while this reduces the potential for sequence errors to influence results, it is possible that low abundance OTUs could also be biologically relevant but would not be taken into account using this estimator. Therefore, I decided to also calculate the Chao1 richness estimator which is calculated using the number of OTUs with only one (singletons) or two reads (doubletons) associated with them (197).

2.15.2 OTUs: Beta-diversity

All tests were carried out within mothur.

Distance matrix creation

Lower-triangle distance matrices describing the similarity of samples by their bacterial compositions were created in order to perform AMOVAs (analysis of molecular variance) and to construct PCOA (principle coordinate analysis) graphs. Distance matrices were created using Yue and Clayton theta values (198) which take into account both the number and abundance of OTUs.

AMOVA (analysis of molecular variance)

AMOVA is a non-parametric analogue of the analysis of variance which analyses whether two groups of samples cluster significantly separately from each other using a distance matrix (199). This test was chosen as when comparing two groups of samples which have the same amount of diversity but different centroids, AMOVA is better at identifying the significantly separate clustering of these groups in comparison to other commonly used tests such as weighted and unweighted UniFrac (200).

Clustering of samples into metacommunities

It is possible to identify whether samples cluster into separate microbial community types without supplying any meta-information. This can be useful when attempting to identify clustering of samples by an unknown variable. The get.communitytype command in mothur clusters samples into metacommunities by their microbial communities using Dirichlet multinomial mixtures (201). This technique models the data using increasing numbers of metacommunities and the model which is found to have the lowest Laplace approximation value is assumed to best fit the data. This test is used in Chapter 4 of this thesis.

Identifying significantly different OTUs between groups

The Metastats package within mothur (202) was used to identify OTUs which were significantly different between groups. This is a non-parametric version of the T-test which controls for multiple statistical tests by employing a false discovery rate. For example, with a P-value = 0.05 you would expect 5 false positives if testing 100 OTUs. Metatstats assumes that the P-values of actually null tests are normally distributed and uses this assumption to test for true positives (q-value <0.05 is truly positive). The indicator command within mothur can also be used to test for OTUs which are more associated with particular groups but unlike metastats it can be used to compare more than two groups (203).

2.15.3 Parametric and non-parametric statistics

When comparing the richness and diversity estimate values of groups there are various tests which can be used depending upon whether the data is parametric or non-parametric and whether the samples are independent or related. Prior to statistical analysis, data were tested for normality both by the construction of histograms and by performing the Shapiro-Wilk test (SPSS) or the Anderson-Darling test (Minitab). For parametric data, I used the standard T-test for independent groups and the paired T-test for related groups. For non-parametric data, I used the Mann-Whitney U or Kruskal-Wallis tests for independent groups and the Wilcoxon Signed Rank test or Friedman's test for related groups. Tests were carried out within either Minitab or SPSS.

2.15.4 Graphs

The relationships between samples based upon the bacterial OTUs they contain are often highly complex and are usually best modelled in dimensions of over three. As such, it is usually impossible to completely accurately represent the relationships between samples visually. PCOA graphs attempt to reduce this complexity as much as possible into 2 or 3 dimension. Each dimension is then assigned a value depicting the percentage of variability it depicts which can be used to calculate how much of the total variability is represented on the graph. PCOA axis coordinates and axis variability percentage values were calculated in mothur and graphs were constructed within Excel or SPSS. Ellipses were added using Inkscape (www.inkscape.org). To identify which OTUs are responsible for moving sample points along the PCOA axes in a specific direction I ran the corr.axes command in mothur which uses the Spearman's rank correlation coefficient to correlate OTUs to axes.

Another way of visualising sample clustering by OTU composition is by the use of heatmaps. Heatmaps were constructed in R Version 3.2.2 (204) using the gplots (205), heatplus (206), RColorBrewer (207) and Vegan (208) packages. Clustering by OTU composition within heatmaps was performed using the Bray-Curtis dissimilarity (209). Other simple graphs such as box-plots and pie-charts were constructed in either Minitab, SPSS or Excel.

Chapter 3: Variability of the Sheep Lung Microbiota

3.1 Introduction and aims

While the use of protected specimen brushings reduces the chance that samples will be contaminated by upper respiratory tract bacteria in comparison to those collected by bronchoalveolar lavage, only a small area of the lung epithelium is able to be sampled using this method. If there is a large amount of variation in the types of microbes found across the lung then it may be difficult to acquire a truly representative set of samples without the number of brushings becoming prohibitively large.

This paper seeks to address the extent of inter- and intra-individual variation present in the sheep lung microbiota. In this study, all work was performed by myself except as further specified: anaesthesia was induced and monitored by Steven Wright, David Collie and Peter Tennant. Bronchoscopic procedures were carried out by Peter Tennant. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording. Sheep showed no clinically overt signs of respiratory illness during this study. A mock community which had been diluted 1:100 was included in this study so that the effect of a low PCR template concentration on PCR bias could be assessed.

3.2 Research paper

This research was published as 'Variability of the Sheep Lung Microbiota' in Applied and Environmental Microbiology (1). For papers published in American Society of Microbiology journals, the American Society of Microbiology does not require authors to obtain permission to include these papers in their thesis, provided the original work is properly cited. Supplemental figures and tables can be found in **Appendix 2**. Data Set S1 can be found as Additional file 3_1.xlsx. Data Set S2 can be found as Additional file 3_2.xlsx.

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology

Variability of the Sheep Lung Microbiota

Laura Glendinning,[a] Steven Wright,[a] Jolinda Pollock,[a,b] Peter Tennant,[a] David Collie,[a] Gerry McLachlan[a]

The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, Midlothian, United Kingdom[a]; Monogastric Science Research Centre, Scotland's Rural College (SRUC), Edinburgh, Midlothian, United Kingdom[b]

ABSTRACT

Sequencing technologies have recently facilitated the characterization of bacterial communities present in lungs during health and disease. However, there is currently a dearth of information concerning the variability of such data in health both between and within subjects. This study seeks to examine such variability using healthy adult sheep as our model system. Protected specimen brush samples were collected from three spatially disparate segmental bronchi of six adult sheep (age, 20 months) on three occasions (day 0, 1 month, and 3 months). To further explore the spatial variability of the microbiotas, more-extensive brushing samples ($n = 16$) and a throat swab were taken from a separate sheep. The V2 and V3 hypervariable regions of the bacterial 16S rRNA genes were amplified and sequenced via Illumina MiSeq. DNA sequences were analyzed using the mothur software package. Quantitative PCR was performed to quantify total bacterial DNA. Some sheep lungs contained dramatically different bacterial communities at different sampling sites, whereas in others, airway microbiotas appeared similar across the lung. In our spatial variability study, we observed clustering related to the depth within the lung from which samples were taken. Lung depth refers to increasing distance from the glottis, progressing in a caudal direction. We conclude that both host influence and local factors have impacts on the composition of the sheep lung microbiota.

IMPORTANCE

Until recently, it was assumed that the lungs were a sterile environment which was colonized by microbes only during disease. However, recent studies using sequencing technologies have found that there is a small population of bacteria which exists in the lung during health, referred to as the "lung microbiota." In this study, we characterize the variability of the lung microbiotas of healthy sheep. Sheep not only are economically important animals but also are often used as large animal models of human respiratory disease. We conclude that, while host influence does play a role in dictating the types of microbes which colonize the airways, it is clear that local factors also play an important role in this regard. Understanding the nature and influence of these factors will be key to understanding the variability in, and functional relevance of, the lung microbiota.

Within the past 5 years, a diverse array of bacteria has been detected in healthy lungs through the use of non-culture-based methods (1, 2). These bacterial communities are commonly referred to as the lung microbiota and are thought to originate predominantly from the upper respiratory tract (3, 4). The presence of particular bacterial communities in the lung has been associated with several human diseases, including cystic fibrosis (5), chronic obstructive pulmonary disease (6), bronchiectasis (7), and lung transplant rejection (8).

While variation in the microbial communities present in the human lung exists at both large and small scales, based upon the location of the bacteria within the lungs (9) and the host cell types present (10), intraindividual variation has been found to be significantly less than interindividual variation, indicating that each individual may play host to a specific lung microbiota (9).

The lung microbiota of healthy domestic sheep has previously been investigated using culture-based methods (11–14), but these studies have shown conflicting descriptions of the extent of lung colonization by bacteria. A study of pneumonic Bighorn sheep lungs found that, for most sheep studied, bacterial 16S rRNA gene amplification and sequencing was able to identify additional bacterial species which were not found by culturing (15). Previous studies have also examined the upper respiratory tracts of healthy sheep by culture-based methods (11, 12, 14, 16). These studies are highly varied in the types and proportions of microbes identified.

Previously, our group studied the composition of the lung microbiota in sheep pre- and postinfection with *Pseudomonas aeruginosa* (17). That study included the first description of the lung microbiota communities of healthy domestic sheep by next-generation sequencing. A diverse community of microbes was identified, and variability was seen to be high, both within and between animals. The variability of the healthy lung microbiotas at specific lung sites over time has not been reported for any animal, although serial sampling of nondiseased humans is planned as part of the Lung HIV Microbiome Project (LHMP) (18).

In the present study, protected specimen brush samples were collected from three spatially disparate segmental bronchi at three time points (baseline, 1 month, and 3 months) to examine the compositions and variability of the lung microbiotas in healthy domestic sheep. In addition, samples were also taken from a sep-

Received 19 February 2016 **Accepted** 15 March 2016

Accepted manuscript posted online 18 March 2016

Citation Glendinning L, Wright S, Pollock J, Tennant P, Collie D, McLachlan G. 2016. Variability of the sheep lung microbiota. Appl Environ Microbiol 82:3225–3238. doi:10.1128/AEM.00540-16.

Editor: P. D. Schloss, University of Michigan

Address correspondence to Laura Glendinning, laura.glendinning@roslin.ed.ac.uk.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.00540-16.

TABLE 1 Sheep used in this study

Sheep ID	Gender	Mean wt (kg) ± SD	Mean rectal temp (°C) ± SD
2D618	Female	51 ± 3.1	39.0 ± 0.06
2S066	Male (castrated)	69 ± 2.6	39.6 ± 0.20
2D619	Female	59 ± 1.7	39.3 ± 0.20
2D620	Female	64 ± 4.6	39.1 ± 0.21
2D644	Female	65 ± 1.0	39.3 ± 0.06
2D645	Female	70 ± 2.0	39.4 ± 0.06

arate sheep from a greater number of respiratory tract locations to further explore the extent of spatial variability.

Such studies are fundamental to understanding the functional relevance of lung microbiota in health and disease in ruminants. Indeed, bacterial pneumonia is well recognized in cattle and sheep and is often associated with high morbidity and mortality. Notably, regional predilection is evident in that infection by *Pasteurella* occurs most frequently in the apical and cardiac lobes in both sheep (12, 19) and cattle (20, 21). Coinfections with other respiratory pathogens are commonplace; it is already well known that infection by *Bordetella parapertussis* and *Mycoplasma ovipneumoniae* can lead to more-severe disease caused by *Mannheimia* (*Pasteurella*) *haemolytica* (22–25), and there are well-recognized links to stressful events, such as housing or transport. As it is conceivable that changes in the lung microbiota may precipitate or associate with such events, it is vital to ground future disease-related studies on a firm basis of understanding normal variation in health. While the immediate focus of such studies relates to animal health, it is also important to acknowledge that sheep are frequently used as models for human respiratory research (26, 27) and that there is an ongoing need to highlight any comparative contrasts and consistencies as and when they arise.

MATERIALS AND METHODS

Animals and airway sampling. Six 20-month-old Suffolk cross sheep were used in this study (Table 1) (5 females, 1 castrated male) and were housed indoors in pens for the trial duration. No animals had undergone bronchoscopic examination during the 4 months preceding the study. Animal procedures were subject to the Animals (Scientific Procedures) Act of 1986 and were approved by the Roslin Institute Animal Welfare and Ethics Committee.

Anesthesia was performed as described previously (28). Sheep were sampled by protected specimen brushings (disposable microbiology brush; ConMed, New York, NY, USA) at 0 days (baseline), 1 month, and 3 months. Sampling sites are shown in Fig. 1. Bronchoscopy was performed via an endotracheal tube by the same operator for all sheep at all time points. The sample harvest dates can be found in Table S1 in the supplemental material. Before sampling of every sheep on any given day, 7.5 ml of phosphate-buffered saline (PBS) was passed through the bronchoscope channel to act as an environmental quantitative PCR (qPCR) control. Bronchoscope washings were centrifuged at 13,000 × *g* for 15 min, and the pellet was resuspended in 500 μl of PBS.

A throat swab and brushing samples (harvested as described above) were also taken from another sheep (female; age, 36 months; 60-kg body weight) at a single time point to further explore the spatial variability of the lung microbiotas (sampling date 1 May 2015). Brushing sites were dorsal and ventral trachea and paired sites from either side of airway bifurcations progressing along the anterior-to-caudal lung axis (Fig. 2).

DNA extraction, amplification, and sequencing. DNA extraction was performed using the Mo Bio (Carlsbad, CA, USA) PowerSoil DNA isolation kit. Brushes were transferred into PowerSoil bead tubes with PowerSoil solution C1 and PowerSoil bead solution. Bead tubes were heated at 65°C for 10 min and then placed in a FastPrep FP120 cell disrupter (Qbiogene, Inc., France) for 45 s at 5.0 m/s. From this point onward, the manufacturer's instructions were followed, except for the final elution step. Purified DNA was eluted into 50 μl rather than 100 μl of PowerSoil solution C6 to increase the DNA concentration.

All PCR steps used Q5 high-fidelity 2× master mix (New England BioLabs, Beverly, MA, USA). A nested PCR was performed with Illumina adaptor sequences and barcodes (see Table S2 in the supplemental material) included only on the primers for the second round in an attempt to reduce bias caused by barcoded primers when amplifying low-biomass samples (29). The conditions for the first round of PCR, amplifying the V1-to-V4 16S hypervariable regions (primers 28F [5′-GAGTTTGATCNTGGCTCAG-3′] and 805R [5′-GACTACCAGGGTATCTAATC-3′]), were as follows: 94°C for 2 min, followed by 20 cycles of 94°C for 1 min, 55°C for 45 s, and 72°C for 1.5 min, followed by 72°C for 20 min. The conditions for the second round of PCR, amplifying the V2-to-V3 16S hypervariable regions (primers 104F [5′-GGCGVACGGGTGAGTAA-3′] and 519R [5′-GTNTTACNGCGGCKGCTG-3′]), were as follows: 98°C for 30 s, followed by 20 cycles of 98°C for 10 s, 67°C for 30 s, and 72°C for 10 s, followed by 72°C for 2 min. Amplicons from both rounds of PCR were purified using the AMPure XP PCR purification system (Beckman Coulter, La Brea, CA, USA). Amplicons were sequenced using an Illumina MiSeq or HiSeq (Illumina, San Diego, CA) run producing paired-end 250-nucleotide reads (30). Those samples sequenced by two MiSeq runs are listed in Data Set S1 in the supplemental material, and those sequenced by HiSeq are listed in Data Set S2. When samples from the MiSeq runs were found to have low read numbers, they were sequenced again on a separate MiSeq run (samples 2D618 RA [right apical] at 3 months and 2D619 RA at 3 months). We previously confirmed cross-run stability by comparing separate runs made on the same samples (Fig. S1).

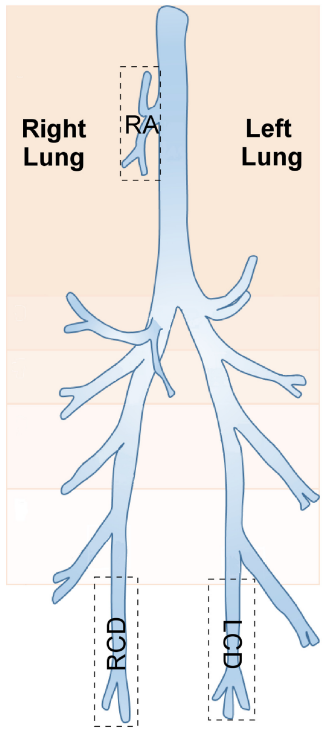
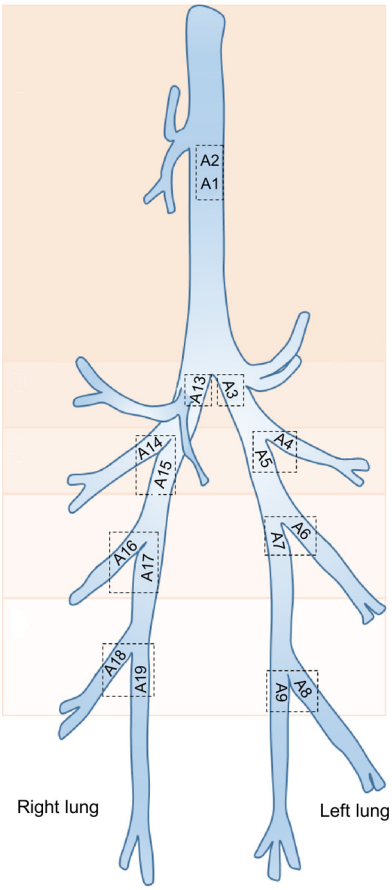


FIG 1 Diagram of a sheep lung, divided into anatomical segments. Boxes indicate the segments where protected specimen brushings were performed in the lungs of six sheep at three time points; these correspond to the right apical (RA), right caudal diaphragmatic (RCD), and left caudal diaphragmatic (LCD) segments.



A1: Ventral trachea just caudal to bifurcation with right apical (RA) lobe segmental bronchus

A2: Dorsal trachea just caudal to bifurcation with right apical (RA) lobe segmental bronchus

A3: Medial aspect of left main bronchus immediately caudal to carina

A4: Medial aspect of first left ventral diaphragmatic segmental bronchus (LVD1), just caudal to bifurcation from left main bronchus

A5: Left main bronchus, lateral wall just after bifurcation to LVD1

A6: Medial aspect of second left ventral diaphragmatic segmental bronchus (LVD2), just caudal to bifurcation from left main bronchus

A7: Left main bronchus, lateral wall just after bifurcation to LVD2

A8: Third left ventral diaphragmatic segmental bronchus (LVD3)

A9: Left caudal diaphragmatic (LCD) segmental bronchus

A13: Medial aspect of right main bronchus immediately caudal to carina

A14: Medial aspect of first right ventral diaphragmatic segmental bronchus (RVD1), just caudal to bifurcation from right main bronchus

A15: Right main bronchus, lateral wall just after bifurcation to RVD1

A16: Medial aspect of second right ventral diaphragmatic segmental bronchus (RVD2), just caudal to bifurcation from right main bronchus

A17: Right main bronchus, lateral wall just after bifurcation to RVD2

A18: Third right ventral diaphragmatic segmental bronchus (RVD3)

A19: Right caudal diaphragmatic (RCD) segmental bronchus

FIG 2 Locations of brushings within sheep lungs. Protected specimen brushings were performed in the sections of the lung labeled A1 to A9 and A13 to A19 in one sheep at one time point.

Extraction kit controls were produced by carrying out a reagent-only extraction using the Mo Bio PowerSoil DNA isolation kit. PCR reagent controls were constructed by adding 20 μl of nuclease-free water to the PCR mixture. The Human Microbiome Project Mock Community HM-782D (100,000 copies per organism per μl; BEI Resources, ATCC, Manassas, VA, USA), extraction kit controls, and PCR reagent-only controls and positive controls (DNA extracted from *Pseudomonas aeruginosa* strain PA0579) were amplified and sequenced by the same methods as were used for samples.

A separate mock community sample was sequenced using an Illumina HiSeq. For this sample, the solution produced from the first round of PCR was diluted 1:100 in nuclease-free water before being used in the second round of PCR. This was carried out to ascertain the effect on PCR bias of placing different concentrations of DNA into the second PCR round.

Bioinformatic and statistical analysis. Primers were removed using Cutadapt (31). Sequences which contained more than one base error per 10 primer bases were removed from further analysis. The following steps were carried out in mothur (32) and were based upon a protocol developed for MiSeq by the mothur creators (30). Forward and reverse reads were aligned to form one continuous DNA sequence; any sequences which failed to align were discarded. Sequences which contained ambiguous bases, were less than 369 bp in length, or contained homopolymers of greater than 9 bp were also discarded. Chimeras were identified and removed using UCHIME (33). Sequences were aligned to the SILVA reference alignment (34) and were classified using mothur's Bayesian classifier against the Greengenes database (35), which was trimmed to the V2-V3 hypervariable region of the 16S rRNA gene to improve classification depth (36). Sequences identified as not originating from bacteria were removed from further analysis. Operational taxonomic units (OTUs) were clustered into phylotypes using a database-dependent approach and then subsampled.

Distance matrices were created using Yue and Clayton theta values (37). Analysis of molecular variance (AMOVA) (38) was used to determine significant differences between the bacterial compositions of groups. Principal-coordinate analysis (PCoA) graphs were constructed to visualize similarities between samples. The inverse Simpson index was used to quantify diversity. Where data were nonparametric, the Friedman test was used to identify significant differences in diversity, using Minitab 16 for Windows (Minitab, Coventry, United Kingdom). All other statistical tests were carried out within mothur. Metastats (39) was used to identify OTUs which were different between groups. Good's coverage (40) was used to estimate sample coverage, and the Chao 1 index was used to calculate richness. Indicator OTUs (OTUs which are indicative of a particular group of samples) were identified using the indicator metric within mothur (41). Repeated-measures analyses of variance (ANOVAs) were carried out using the Vegan package in R (42–44).

qPCR. qPCRs were performed using the LightCycler 480 SYBR green I master mix (Roche Applied Science, Indianapolis, IN, USA), 1 μl of extracted DNA solution, and the 16S rRNA gene qPCR primers UniF340 (5′-ACTCCTACGGGAGGCAGCAGT-3′) and UniR514 (5′-ATTACCG CGGCTGCTGGC-3′) at a final concentration of 0.4 μM.

The qPCR run consisted of a preincubation step of 50°C (ramp rate, 4.80°C/s for 2 min) and then 95°C (ramp rate, 4.80°C/s for 10 s) and an amplification step consisting of 45 cycles of 95°C (ramp rate, 4.80°C/s for

30 s) and then 63°C (ramp rate, 2.50°C/s for 30 s). This was followed by a melting cycle consisting of 95°C (ramp rate, 4.80°C/s for 5 s) and then 65°C (ramp rate, 4.80°C/s for 1 min), followed by 97°C (ramp rate, 0.11°C/s; acquisition mode, continuous).

Negative controls consisted of both water and extraction kit reagent controls. For water controls, 1 μl of nuclease-free water was added to the qPCR mixture. For extraction kit controls, DNA extractions were carried out using the Mo Bio PowerSoil DNA isolation kit (Carlsbad, CA, USA) by following the same protocol as was used to extract DNA from samples, except that no sample was added, meaning that any bacterial DNA in the final elution must have been derived from the extraction kit reagents. Then, 1 μl of this elution was added to the qPCR mixture.

In order for us to compare the quantities of bacterial DNA found in bronchoscope wash and brushing samples, it was necessary to use a unit of measurement which could be applied to both sample types. Bacterial DNA concentrations are therefore reported as the 16S copy numbers present per microliter of eluent produced from samples by the Mo Bio PowerSoil DNA isolation kit. Statistical analysis was carried out in Minitab 16 for Windows. When data were nonparametric, the Mann-Whitney U test was used to statistically compare groups.

Nucleotide sequence accession numbers. The unassembled reads, with primers removed, are publicly available through the NCBI Sequence Read Archive (SRA) under the BioProject accession no. PRJNA298882.

RESULTS

Quality control and adequacy of sequencing. After DNA sequences were constructed from the forward and reverse reads generated by sequencing, various quality control steps were performed to decrease the number of artifacts and poor-quality sequences used in subsequent analyses.

For the MiSeq runs, these steps resulted in a 15% loss of sequences (sequencing error rate = 0.39%). On average, samples contained 205,625 ± 27,232 (mean ± standard error of the mean [SEM]) sequences and a total of 925 bacterial OTUs were identified (see Data Set S1 in the supplemental material). Sequences were assigned to OTUs based on their taxonomic classifications. Each OTU does not necessarily represent an individual bacterial species but instead represents the lowest taxonomic level to which its bacterial sequences could be assigned. For example, 77.4% of reads could be identified to the genus level, while 31.1% could be assigned to the species level. If two species from the same genus could be assigned only to the genus level, then both were binned into the same OTU.

For the HiSeq run, samples contained on average 233,505 ± 69,735 (mean ± SEM) sequences, and the sequencing error rate was 0.39%. Six hundred thirty-three OTUs were identified (see Data Set S2 in the supplemental material), and the total reduction in sequence numbers due to quality control was 5%.

Good's coverage estimate values exceeded 97% for all samples. This indicates that at least 97% of the bacteria present in our original samples were likely to have been identified, demonstrating that the depth of sequencing was adequate.

Of the 20 bacteria contained in the mock community, all could be taxonomically identified down to genus level, except that *Bacillus cereus*, *Escherichia coli*, and *Listeria monocytogenes* could be identified only to the family level. This indicates that the primers were able to amplify a wide diversity of bacteria. While the proportions of bacterial DNA were different from the proportions anticipated if no PCR bias was present (Table 2), this was less apparent in the sample which had been diluted 1:100 after the first round of PCR. In the undiluted mock community, the proportions of bacterial orders differed from the expected proportions by an average of 9.48% (SEM, 2.24%; range, 0.99% to 19.48%), whereas the orders in the diluted mock community differed on average by 4.33% (SEM, 1.12%; range, 0.29% to 12.71%). This diluted mock community may be more comparable to the kind of biases we found in our samples, as the undiluted mock community contained a far higher concentration of template DNA (2,000,000 16S copies per μl) than our samples did on average (13,133 16S copies per μl).

We assumed that PCR bias could reasonably be expected to apply equally across all samples and, therefore, that any statistical tests between samples should still be valid. The two bacterial species most overrepresented in the undiluted mock community (*Deinococcus radiodurans* and *Helicobacter pylori*) are not commonly associated with the respiratory tract, and bacteria from these genera were very rare within our data set.

Longitudinal study in 6 sheep over 3 months. To examine the spatial, longitudinal, and interindividual variations of the sheep lung microbiota, lung brushing samples were taken from 3 spatially disparate lung locations (right apical [RA], right caudal diaphragmatic [RCD], and left caudal diaphragmatic [LCD]) in 6 sheep at 3 time points (baseline, 1 month, and 3 months). Estimates of total bacterial yield from qPCR analysis indicated that sheep lung brushing samples contained an average of 13,133 ± 894 (mean ± SEM) 16S copy numbers/μl (range, 1,032 to 37,627 16S copy numbers/μl). Bronchoscope wash control samples contained significantly lower bacterial 16S rRNA gene concentrations than lung brushing samples (Mann-Whitney U test, $P < 0.0001$), containing an average of 1,471 ± 279 (mean ± SEM) 16S copy numbers/μl (range, 397 to 4,792 16S copy numbers/μl) (Fig. 3). The qPCR-negative water controls were found to contain 190, 479, and 739 16S copy numbers/μl, and the extraction kit controls were found to contain 347 and 511 16S copy numbers/μl.

After sequencing and subsampling, bacterial communities isolated from the extraction kit and 16S PCR-negative controls were found to cluster separately from those found in sheep lung brushing samples (AMOVA, $P < 0.001$). Extraction kit controls were included from two different lots. The most abundant OTUs found in the first extraction kit control were *Corynebacterium* (36%), *Enterobacteriaceae* (13%), *Mycobacterium llatzerense* (7%), and *Staphylococcus haemolyticus* (5%). The most predominant OTUs in the second extraction kit control were *Aerococcus* (13%), *Dermabacteraceae* (11%), *Micrococcus* (10%), *Enhydrobacter* (9%), and *Leuconostoc* (7.2%). The predominant bacterial order present in both extraction kit controls was *Actinomycetales* (50.1% and 40.5%, respectively).

The bacteria isolated from lung brushing samples predominantly belonged to the orders *Bacillales* (26%), *Actinomycetales* (21%), *Clostridiales* (11%) and *Lactobacillales* (9%), while common genera included *Staphylococcus* (16%), *Corynebacterium* (9%), *Jeotgalicoccus* (5%), and *Streptococcus* (5%).

The underlying changes in bacterial OTUs between sampling points were examined. The bacterial communities found in lung brushing samples clustered significantly by time point (AMOVA, $P < 0.001$) (Fig. 4). The OTUs causing this clustering were identified by applying Metastats (see Tables S3 and S4 in the supplemental material). The largest difference observed between the first and second time points was an 11% increase in the abundance of an OTU identified as *Corynebacterium*. This is also the most abundant OTU in one of our extraction kit controls. OTU 12, *Mycobacterium llatzerense*, was also significantly more abundant at the

TABLE 2 Proportions of DNA sequence reads belonging to bacterial members of a mock community

Taxonomy (order or genus)	Expected proportion of reads (%)	Actual proportion of reads (%)		Mock community species
		Undiluted	1:100 dilution	
Order				
Deinococcales	5	24.48	7.65	*Deinococcus radiodurans*
Campylobacterales	5	22.05	12.65	*Helicobacter pylori*
Bacteroidales	5	19.59	10.91	*Bacteroides vulgatus*
Bacillales	20	8.60	22.40	*Bacillus cereus*, *Listeria monocytogenes*, *Staphylococcus aureus*, *Staphylococcus epidermidis*
Lactobacillales	25	5.10	12.29	*Enterococcus faecalis*, *Lactobacillus gasseri*, *Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pneumoniae*
Clostridiales	5	4.01	7.86	*Clostridium beijerinckii*
Rhodobacterales	5	3.92	5.29	*Rhodobacter sphaeroides*
Pseudomonadales	10	3.42	5.97	*Acinetobacter baumannii*, *Pseudomonas aeruginosa*
Enterobacteriales	5	3.33	5.52	*Escherichia coli*
Neisseriales	5	2.17	3.49	*Neisseria meningitidis*
Actinomycetales	10	1.27	2.92	*Actinomyces odontolyticus*, *Propionibacterium acnes*
Other/unclassified	0	2.03	3.08	
Genus[a]				
Deinococcus	5	24.33	7.61	*Deinococcus radiodurans*
Helicobacter	5	22.04	12.65	*Helicobacter pylori*
Bacteroides	5	19.59	10.90	*Bacteroides vulgatus*
Rhodobacter	5	3.91	5.29	*Rhodobacter sphaeroides*
Clostridium	5	3.73	7.59	*Clostridium beijerinckii*
Staphylococcus	10	3.04	7.58	*Staphylococcus aureus*, *Staphylococcus epidermidis*
Lactobacillus	5	2.77	6.59	*Lactobacillus gasseri*
Pseudomonas	5	2.33	3.70	*Pseudomonas aeruginosa*
Neisseria	5	2.15	3.27	*Neisseria meningitidis*
Enterococcus	5	1.40	2.63	*Enterococcus faecalis*
Acinetobacter	5	0.97	1.62	*Acinetobacter baumannii*
Propionibacterium	5	0.76	1.77	*Propionibacterium acnes*
Actinomyces	5	0.48	1.12	*Actinomyces odontolyticus*
Streptococcus	15	0.47	1.63	*Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pneumoniae*
Other/unclassified	0	12.03	26.05	

[a] The species *Bacillus cereus*, *Escherichia coli*, and *Listeria monocytogenes* could not be classified to the genus level.

1-month time point and was the third-most-abundant OTU in the same extraction kit control. It therefore is likely that our time points were affected to different degrees by reagent contamination and that the analysis of segments over time is not possible. However, all samples taken in the same sheep at the same time point were processed using the same extraction kit; therefore, an analysis of spatial variability could be performed.

Visual perceptions of community structure indicated that, in some sheep, samples taken from separate lung sites differed appreciably, whereas in other sheep, there appeared relative concordance between such samples (see the example shown in Fig. 5). A full visual summary of the results can be found in Fig. S2 in the supplemental material. There were no significant differences between the levels of diversity of communities located at different lung sites (inverse Simpson index, Friedman test, $P > 0.5$).

Sheep clustered separately by the compositions of their lung bacterial communities at the baseline time point (AMOVA, $P = 0.001$) and at the 3-month time point (AMOVA, $P = 0.045$), indicating that samples taken from within the same sheep were more similar to one another than to samples taken from other sheep. At the 1-month time point, sheep did not cluster in this manner (AMOVA, $P = 0.394$), though this is likely due to the presence of contamination causing a homogenization of our 1-month samples. Pairwise comparisons of samples showed no significant results. The similarity of samples to one another can be visualized using PCoA graphs (Fig. 6).

Spatial variability of the lung microbiota in an individual sheep. The observed variability between spatially disparate lung sites in some sheep prompted enquiry as to the consistency of bacterial communities sampled from sites in close spatial apposition.

Further samples were derived by systematically sampling multiple sites of the lungs of an individual animal at one time point. While the 3-month experiment did not include a control for every lot of extraction kit used, emerging literature and opinion within the field have since indicated the value of using the same extraction kit for all samples. This strategy, therefore, was adopted for these latter samples, which were all processed at the same time.

The extraction kit control was mainly composed of one OTU (OTU 18: 79%), which was also present in our brushing samples (mean ± SEM, 51.1% ± 3.3%). We felt confident in removing this OTU from all of our samples prior to analysis, as it could be identified to the species level (*Methylobacterium komagatae*) and was considered highly unlikely to be found within the sheep lung. No further OTUs were removed before analysis.

Lung brushing samples contained on average 2,116 16S copy

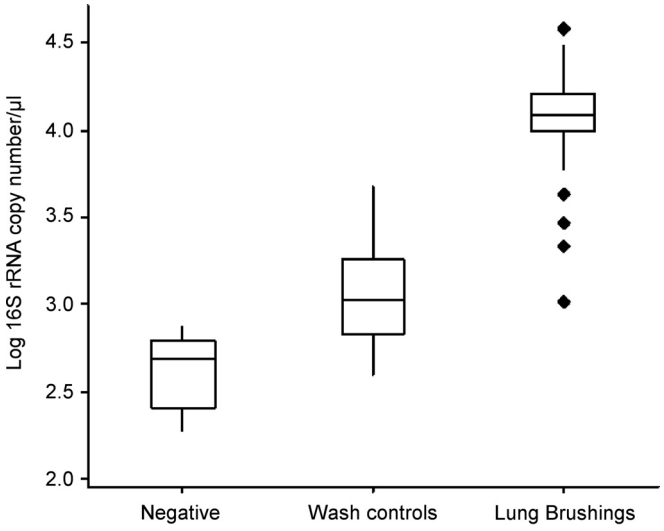


FIG 3 qPCR of lung brushing and control specimens. The bronchoscope channel was flushed with 7.5 ml of PBS, and the wash was collected (wash control, $n = 18$) prior to protected specimen brushing being performed on the lungs of sheep (lung brushings, $n = 54$). DNA was extracted from wash control and lung brushing specimens, and the quantity of bacterial DNA was calculated using 16S rRNA gene qPCR. Lung brushing specimens were found to contain significantly higher quantities of bacterial DNA than did wash controls (Mann-Whitney U test, $P < 0.0001$). Negative controls consisted of either water ($n = 3$) or extraction kit ($n = 2$) controls. Boxes, interquartile ranges; diamonds, outliers.

numbers per µl (SEM, 365 copy numbers per microliter), while the throat swab and extraction kit control contained 42,480 and 43 16S copy numbers per microliter, respectively. The richness and diversity of the lung samples (Chao, 103.77 ± 7.32; inverse Simpson index, 14.24 ± 2.14) were found to be far lower than in the throat swab (Chao, 257.038; inverse Simpson index, 9.19). Sample A1, taken from the ventral aspect of the trachea just caudal to the bifurcation with the right apical lobe segmental bronchus, had the second-highest richness (Chao, 155.024) and diversity (inverse Simpson index, 8.713). However, sample A2, which was taken at the same level as sample A1 but from the dorsal aspect of the trachea, had much lower richness (Chao, 76.038) and diversity (inverse Simpson index, 4.925).

The compositions of the communities taken from the respiratory tract showed some variation, even between paired samples located very close to one another (Fig. 7). Subtracheal samples paired to their most proximate neighbor did not cluster together significantly when OTUs were defined at the lowest taxonomic depth (AMOVA, $P = 0.30$). However, paired samples did cluster significantly by the bacterial orders which they contained (AMOVA, $P = 0.046$). Subtracheal samples also clustered significantly (by order) based upon the depth in the lung from which samples were taken (AMOVA, $P = 0.033$) (Fig. 8) (lung depth in this context refers to increasing distance from the glottis, progressing in a caudal direction). An indicator OTU for the group which included the samples A4, A5, A14, and A15 was found to be OTU 4, *Pseudomonadales* ($P = 0.042$). The most abundant bacterial orders identified from brushings were *Clostridiales* (25.8%),

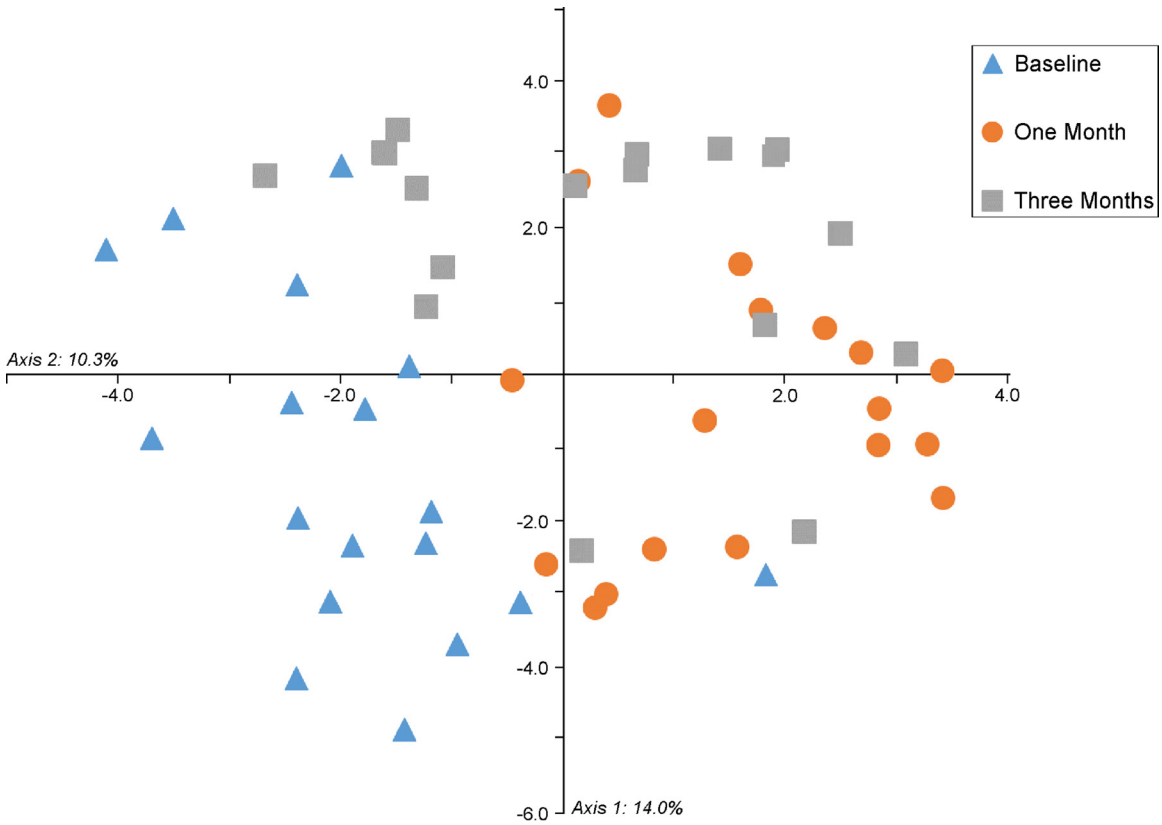


FIG 4 Clustering of time points by lung microbiota composition. A PCoA graph shows the similarities between bacterial communities sampled from three sheep lung segments in six sheep at three time points. Samples were found to cluster significantly by the time point at which they were taken (AMOVA, $P < 0.001$).

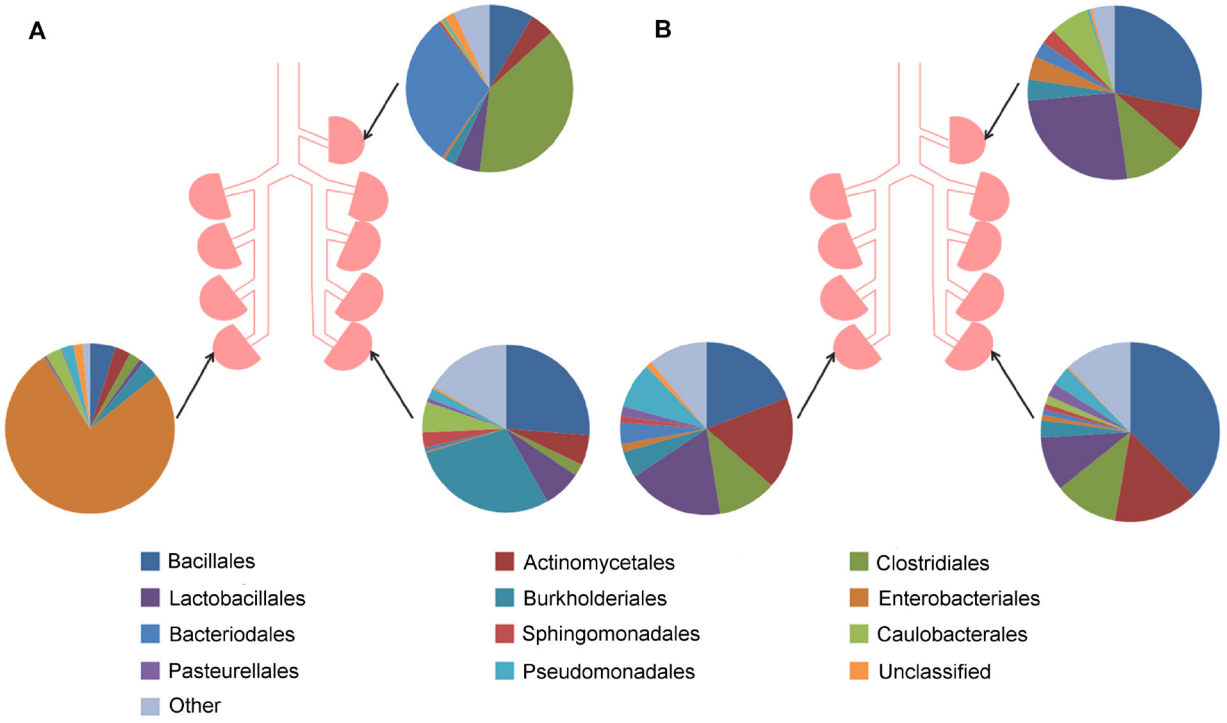


FIG 5 Bacterial communities found in three separate lung segments within two sheep. Protected specimen brushings were performed on the lungs of sheep at three different lung segments (RA, RCD, and LCD) at day 0. Sheep A (2S066) had highly different bacterial communities at each lung segment, whereas sheep B (2D644) had similar bacterial communities at all three lung sites.

Pseudomonadales (18.3%), and *Actinomycetales* (16.0%), while the throat swab was dominated by *Pasteurellales* (36.5%) and *Pseudomonadales* (15.1%). The extraction kit control was predominantly composed of *Actinomycetales* (31.1%) and *Pseudomonadales* (31.0%).

As the *Pasteurellales* order contains several species which are known to act as sheep lung pathogens and which display regional patterns of infection, we felt it would be interesting to investigate where OTUs belonging to this order were found within the respiratory tract (Table 3). By far, the largest proportion of these OTUs was found in the throat swab and in one of the tracheal brushing samples (sample A1).

DISCUSSION

In order to better understand the variability present in the sheep lung microbiotas, we compared the lung bacterial communities of six sheep at three different lung sites over a duration of 3 months. To further explore the extent of spatial variability, we also took 17 samples from the respiratory tract of one sheep.

Previously, the bacteria in healthy domestic sheep lungs had

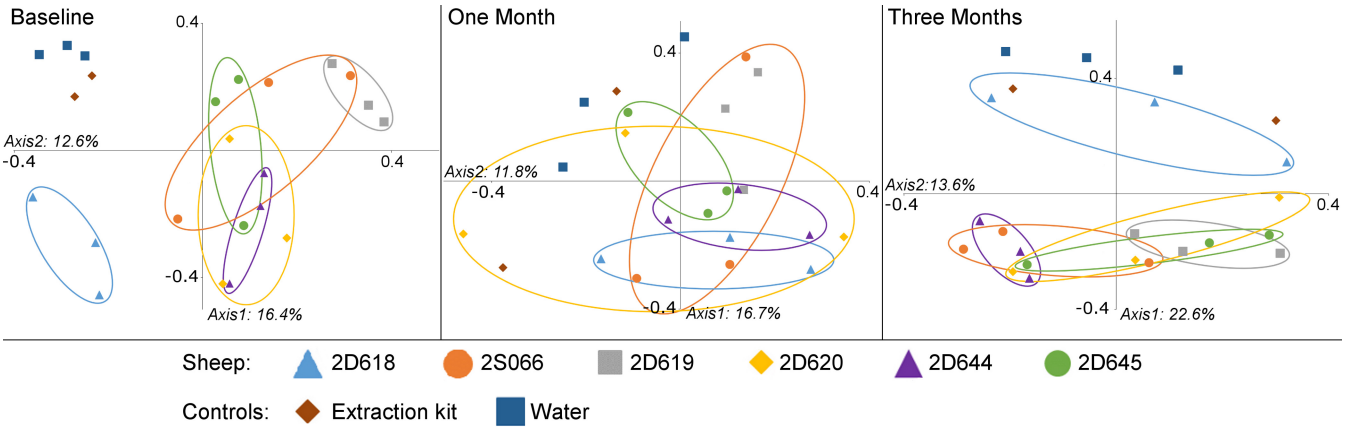


FIG 6 Clustering of individuals by lung microbiota composition. PCoA graphs show the similarities between the bacterial communities extracted from protected specimen brushing samples taken from sheep lungs at three time points (baseline [0 days], 1 month, and 3 months). Samples were taken from three separate lung segments (RA, RCD, and LCD). Samples from within the same sheep were found to cluster significantly at baseline (AMOVA, $P = 0.001$) and at 3 months (AMOVA, $P = 0.045$) but not at 1 month. This is likely to be due to the presence of contaminants originating from the extraction kits in the 1-month samples.

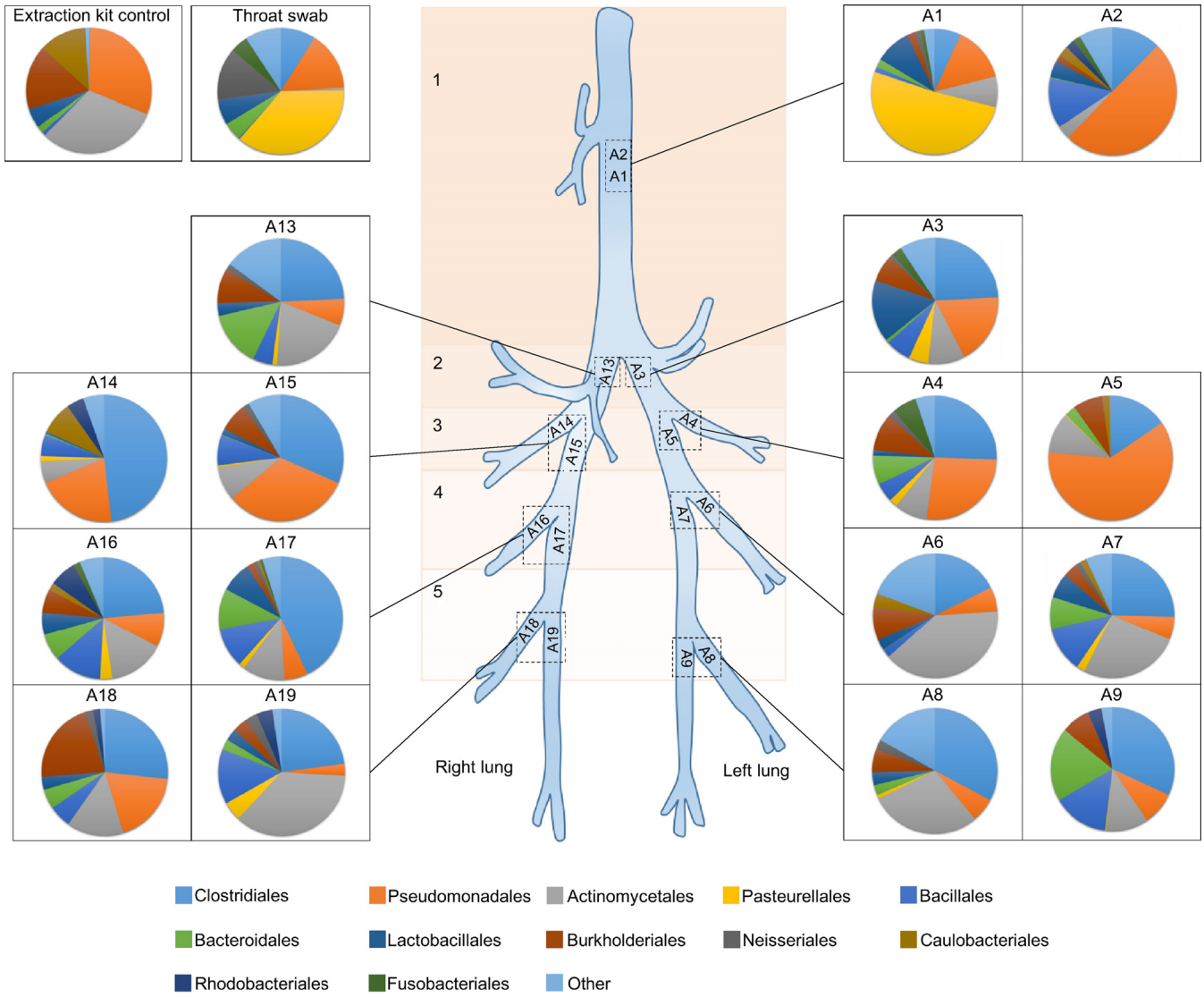


FIG 7 Diagram of the bacterial orders found in the sheep lung. Bacterial orders found in protected specimen brushing samples from the lung and trachea (A1 to A9 and A13 to A19), a throat swab, and an extraction kit control taken during a study of one sheep at one time point.

been investigated by culture-based methods, which seemed to indicate that bacterial colonization of the sheep lung was rare or did not occur in all sheep (11, 12, 14). In contrast, using non-culture-based methods, we have found that all of the sampled sites in our seven sheep harbored diverse communities of bacteria, although in far smaller numbers than is generally found in other niches, such as the gut or upper respiratory tract.

Bacteria belonging to genera previously isolated from goat and sheep lungs (11, 12) were found in our samples. These included *Corynebacterium*, *Bacillus*, *Enterococcus*, *Klebsiella*, *Mannheimia*, *Micrococcus*, *Moraxella*, *Pasteurella*, *Pseudomonas*, *Staphylococcus*, and *Streptococcus*. Of the most common genera observed within our animals, *Staphylococcus*, *Streptococcus*, and *Corynebacterium* are commonly isolated from the upper respiratory tracts and skin of many animals, whereas *Jeotgalicoccus* is a less well-known genus (45) which has not been found to make up a substantial part of the lung microbiota communities in any previous studies. However, it has been isolated from the small intestinal mucosa of calves (46), the canine oral cavity (47), aerosol samples from a poultry house (48, 49), cattle teats (50), lamb meat (51), the rumen of cattle (52), and aerosol samples near a dairy (53).

The most common bacterial orders found in the sheep lung during the 3-month study were *Bacillales*, *Actinomycetales*, and *Clostridiales*. This agrees with the findings of a previous study carried out by our group, which examined the sheep lung microbiota before and after infection with *Pseudomonas aeruginosa* (17). *Pseudomonadales* (mainly *Pseudomonas*) was also commonly found in the lungs during our single-sheep study, while the throat swab from this study was dominated by *Pasteurellales* and *Pseudomonadales*.

Coinfection with *Bordetella parapertussis* or *Mycoplasma ovipneumoniae* has been shown to lead to more-severe disease caused by *Mannheimia* (*Pasteurella*) *haemolytica* (22–25). Mycoplasmas were very rare within our data set, with only one sheep segment containing reads from this genus at one time point. We did not identify any OTUs as *Bordetella*; however, we did find an OTU

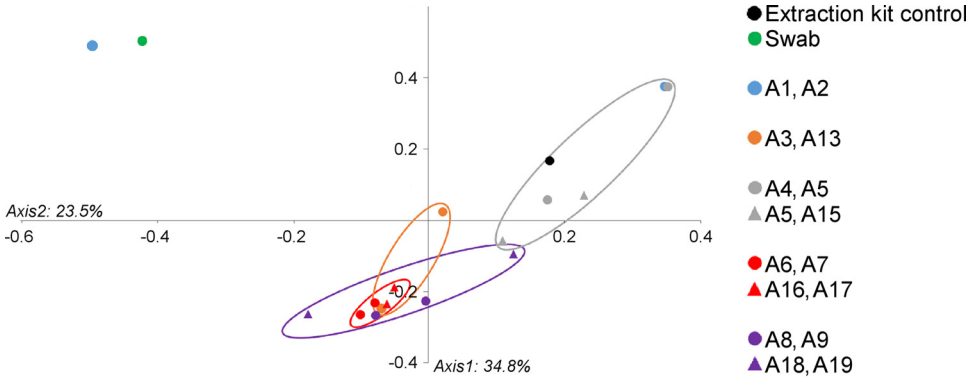


FIG 8 Clustering of lung brushing samples by depth within the lung. A PCoA graph shows the similarity of samples taken at different lung depths based upon the bacterial orders present. Lung depths are represented by color and correspond to different distances from the glottis, progressing in a caudal direction. Adjacent pairs of samples are represented by the same symbol and color. For the exact location of each sampling site, see Fig. 2. Subtracheal samples (≥A3) clustered significantly by lung depth (AMOVA, $P = 0.033$), as did paired samples (AMOVA, $P = 0.046$).

designated *Alcaligenaceae* (the family to which *Bordetella* belongs), though this was uncommon and occurred in low abundance. We identified several OTUs which were classified as members of the *Pasteurellaceae* family, including *Mannheimia*, *Bibersteinia*, and, less commonly, *Aggregatibacter segnis*, *Haemophilus parainfluenzae*, *Bibersteinia trehalosi*, and *Actinobacillus parahaemolyticus*. All of these microbes have previously been isolated from the lungs or upper respiratory tract (54–58). Despite the fact that disease caused by members of this family is often located in the apical and cardiac lobes (12, 19), we observed members of this family to be present across the lung.

The composition of the lung microbiota found in our sheep shows some differences from that previously identified in humans, where *Bacteroidales* are found in higher numbers and there are generally fewer members of the *Actinomycetales* and *Clostridiales* orders (2, 9, 59). Segal et al. identified various bacterial taxa that were commonly found in high relative abundance in human lungs (1). These included taxa which were found in all of our sheep samples in high relative abundance (*Streptococcus*, *Staphylococcus*, *Corynebacterium*), taxa which were found in the majority of our samples but in lower abundances (*Propionibacterium*, *Pseudomonas*), and taxa which were found only sporadically in our samples and were usually in low abundance (*Stenotrophomonas*, *Prevotella*, *Veillonella*, *Fusobacterium*, *Porphyromonas*).

Such differences may at least in part reflect the different surroundings in which sheep live, as well as behavioral or physiological features, such as rumination. A study using buccal swabs to identify bacteria originating from the rumen suggested that, as the time between regurgitation and sampling increases, the orally associated bacterial populations in the buccal cavity increase and the

TABLE 3 Abundances of the OTUs within the *Pasteurellaceae* family found in different locations of the sheep respiratory tract

Specimen type, location, or sample	% of organisms that were in:			
	OTU 5, *Mannheimia*	OTU 6, *Pasteurellaceae*	OTU 7, *Bibersteinia*	OTU 9, *Bibersteinia trehalosi*
Throat swab	23.7	10.1	1.8	0.7
Trachea				
A1	5.5	4.5	28.4	5.3
A2	0	0.03	0	0.01
Left lung				
A3	0	3.04	0.01	0.03
A4	1.2	0.006	0	0
A5	0	0.2	0.006	0
A6	0	0	0	0
A7	0.7	1.4	0.006	0.006
A8	0	0	0.8	0
A9	0.006	0.02	0.2	0
Right lung				
A13	0.9	0	0	0.006
A14	0.006	0.3	0	0
A15	0	0.6	0.006	0.006
A16	2.3	0.6	0.006	0
A17	0.10	0	1.3	0.01
A18	0	0.02	0.01	0
A19	3.2	0	1.3	0

rumen-associated bacteria decrease, potentially contributing to interanimal variation (58). In future studies, it may be useful to take rumen and upper respiratory tract samples alongside lung samples to explore whether the variations between these sites and the lung are related.

Regardless of the highlighted differences between sheep and human lung microbiotas, there is a pressing need to understand the mechanisms that underlie the spatial and temporal variability of microbiota in the mammalian lung. These fundamental studies are difficult to facilitate in healthy human subjects as a consequence of the invasive nature of the repeated sampling protocol as well as the difficulty of controlling for the influence of environmental and/or lifestyle factors. Large-animal models can, however, play an important role in filling this need. Indeed, the physiological and immunological similarities between sheep and human lungs (60, 61) have contributed to the widespread use of sheep as translational models for human lung research (26, 27), including for asthma (62–65), the delivery of drugs via the upper respiratory tract (66–68), emphysema (69–71), pulmonary hypertension (72–74), physical lung injury (75–78), lung infection (28, 79–81), respiratory distress syndromes (82–85), asbestosis (86–88), and lung cancer (89, 90).

In our study, we examined the variability of the lung microbiota in sheep. Bacterial populations were often different between lung segments and between individuals, which confirms our previous observations (17). There was more similarity between samples from the same sheep at the baseline and 3-month time points than between samples taken from different sheep, but this was not found to be the case at the 1-month time point. Lung sample clustering by individuals has previously been identified in humans (9) and sheep (17).

Clearly, large differences can exist in the microbiota sampled from different lung segments at the same time point. This spatial variability of lung microbial populations can be observed in *P. aeruginosa* infections in cystic fibrosis patient lungs (91). The mechanisms underlying such observations have yet to be elucidated; however, possible candidate influences may include regional variability of physiological parameters, such as gas concentrations, osmolality, temperature, pH, and blood flow (92–96), which may lead to the creation of microhabitats providing a selective advantage to certain bacteria (97). It has previously been demonstrated that differences in pH can lead to changes in the colonic microbiota (98) and that temperature combined with humidity can lead to changes in the composition of the skin microbiota (99).

A longitudinal analysis of the lung microbiota at specific lung sites in healthy individuals has not previously been reported. Our goal was to define the variability of the lung microbiota over time and to detect whether there was a sheep lung microbiota "signature" which remains stable. Unfortunately, at the time of carrying out this study, the extent of the variability of bacterial DNA found within different lots of extraction kits was not yet known (100). While we, therefore, did include some extraction kit controls for our longitudinal study, we did not include controls for all lots which were used. Samples from different time points were also processed at different times. Due to our small sample sizes and the fact that samples clustered significantly by time point, we do not feel that accurate conclusions can be drawn about the temporal stability of the microbiota from our data. However, all samples taken from the same time point in the same sheep were processed at the same time. Therefore, we can be confident that the spatial variability that we observed within animals was not due to our methodology.

In some individuals, samples taken from different lung segments were found to be highly different from one another, whereas in others, the lung microbiota appeared to be quite stable across the lung sites. Another finding was the disappearance of the significantly separate clustering of sheep samples at the 1-month time point. This was correlated with an increase in the proportions of several OTUs found in sheep lungs, the most noticeable increase arising from an OTU classified as *Corynebacterium*, which was also the most abundant OTU in samples from one of our extraction kit controls. It is likely that the disappearance of significant clustering by individual at the 1-month time point is due to the increased presence of contamination in our samples.

OTUs that were identified in both samples and negative PCR and extraction kit controls were not removed from the analysis for the 3-month sheep study. The reason for this decision was that a number of bacteria commonly associated with the upper and lower respiratory tracts were present in these controls, including the genera *Streptococcus* and *Pseudomonas*, and it was judged that their removal would merely introduce another source of bias.

Equally, any specific *a priori* manipulation based around assumptions gleaned from the human literature regarding microbiota in the upper and lower respiratory tract are potentially ill advised. Indeed, it has been demonstrated that the microbes found in the lungs of animals often match those found in their bedding and hay (101). It is therefore not possible to dismiss environmental microorganisms as being due only to the contamination of samples.

In our spatial-variation study, one OTU was removed before analysis, as we felt confident that its presence was due to contamination of our extraction kit. Clustering of organisms in lung brushing samples by the lung depth from which they were taken was observed when OTUs were defined by bacterial order. Microorganisms in samples paired with their proximate neighbors were also found to cluster significantly separately from those in brushing samples taken elsewhere in the lung, but this may just be due to the fact that these samples were taken from the same lung depth. Certainly, further research to explore the relationship between lung depth and community composition appears warranted.

After sequencing a mock community of bacteria which contained equimolar concentrations of each bacterial species, we did find some bias present, with some bacterial species being overrepresented or underrepresented. These biases, which may be caused by various factors, including primer mismatching, PCR cycle number, and the bioinformatic pipeline used, are quite common in 16S sequencing (102–105). We also sequenced a 1:100 dilution of the same mock community and found that the apparent biases were far smaller. As the concentration of bacterial DNA in our samples was far lower than that of the undiluted mock community, we feel that the 1:100 dilution is likely to better represent the biases which may be present in our samples, as it is closer to their bacterial DNA concentrations. We believe that this vindicates our choice of DNA amplification strategy, including the use of nested PCR.

It may not be possible to claim that the bacterial abundances identified via 16S sequencing quantitatively represent the relative abundances of bacteria in the sample. Indeed, this is made even more difficult, as different bacterial taxa contain different copy

numbers of the 16S rRNA gene (106). However, it seems logical to assume that, if the same methodology is used for all samples within a study, then the biases present will be the same for all samples, and therefore, comparisons between groups or claims about the types of microbes present in samples would still be valid.

In conclusion, we observed variability in sheep lung microbiotas both between and within individuals. In some animals, different lung segments contained highly different bacterial communities, whereas other animals showed similar communities at all lung sites. While spatial variation was observed to occur over both large and small distances across the lung, samples taken at the same lung depth clustered together separately from those taken at different lung depths. Further studies are needed to explore the stability of the healthy lung microbiota over time.

ACKNOWLEDGMENTS

We thank the administrative staff of the Biotechnology and Biological Sciences Research Council (BBSRC) EASTBIO DTP program for their encouragement and support. We thank the staff at the Dryden farm for providing the everyday care of our animals. We also thank Helen Brown for her advice on statistical methods and sample sizes and Tina Baker for providing the qPCR protocol. The Human Microbiome Project mock community HM-782D was obtained through BEI Resources, NIAID, NIH, as part of the Human Microbiome Project (genomic DNA from microbial mock community B [even, low concentration], v5.1L, for 16S rRNA gene sequencing [HM-782D]).

Funding for this study was supplied by the BBSRC (under the EASTBIO doctoral training program grant BB/J01446X/1). The funders had no input on study design, data collection and interpretation, or the decision to submit this study for publication.

FUNDING INFORMATION

This work, including the efforts of Gerry McLachlan, was funded by Biotechnology and Biological Sciences Research Council (BBSRC) (BB/J01446X/1).

REFERENCES

1. **Segal L, Alekseyenko A, Clemente J, Kulkarni R, Wu B, Chen H, Berger K, Goldring R, Rom W, Blaser M, Weiden M.** 2013. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. Microbiome **1:**19. http://dx.doi.org/10.1186/2049-2618-1-19.
2. **Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG.** 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med **184:**957–963. http://dx.doi.org/10.1164/rccm.201104-0655OC.
3. **Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM.** 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med **187:**1067–1075. http://dx.doi.org/10.1164/rccm.201210-1913OC.
4. **Charlson E, Bittinger K, Chen J, Diamond J, Li H, Collman R, Bushman F.** 2012. Assessing bacterial populations in the lung by replicate analysis of samples from the upper and lower respiratory tracts. PLoS One **7:**e42786. http://dx.doi.org/10.1371/journal.pone.0042786.
5. **Delhaes L, Monchy S, Fréalle E, Hubans C, Salleron J, Leroy S, Prevotat A, Wallet F, Wallaert B, Dei-Cas E, Sime-Ngando T, Chabé M, Viscogliosi E.** 2012. The airway microbiota in cystic fibrosis: a complex fungal and bacterial community: implications for therapeutic management. PLoS One **7:**e36313. http://dx.doi.org/10.1371/journal.pone.0036313.
6. **Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB.** 2011. Analysis of the lung microbiome in the healthy smoker and in COPD. PLoS One **6:**e16384. http://dx.doi.org/10.1371/journal.pone.0016384.
7. **Boyton RJ, Reynolds CJ, Quigley KJ, Altmann DM.** 2013. Immune mechanisms and the impact of the disrupted lung microbiome in chronic bacterial lung infection and bronchiectasis. Clin Exp Immunol **171:**117–123. http://dx.doi.org/10.1111/cei.12003.
8. **Borewicz K, Pragman AA, Kim HB, Hertz M, Wendt C, Isaacson RE.** 2013. Longitudinal analysis of the lung microbiome in lung transplantation. FEMS Microbiol Lett **339:**57–65. http://dx.doi.org/10.1111/1574-6968.12053.
9. **Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Beck JM, Huffnagle GB, Curtis JL.** 2015. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. Ann Am Thorac Soc **12:**821–830. http://dx.doi.org/10.1513/AnnalsATS.201501-029OC.
10. **Dickson R, Erb-Downward J, Prescott H, Martinez F, Curtis J, Lama V, Huffnagle G.** 2014. Cell-associated bacteria in the human lung microbiome. Microbiome **2:**28. http://dx.doi.org/10.1186/2049-2618-2-28.
11. **Yimer N, Asseged B.** 2007. Aerobic bacterial flora of the respiratory tract of healthy sheep slaughtered in Dessie municipal abattoir, northeastern Ethiopia. Rev Med Vet Toulouse **158:**473–478.
12. **Alley MR.** 1975. The bacterial flora of the respiratory tract of normal and pneumonic sheep. N Z Vet J **23:**113–118. http://dx.doi.org/10.1080/00480169.1975.34211.
13. **Alley MR, Quinlan JR, Clarke JK.** 1975. The prevalence of *Mycoplasma ovipneumoniae* and *Mycoplasma arginini* in the respiratory tract of sheep. N Z Vet J **23:**137–141. http://dx.doi.org/10.1080/00480169.1975.34214.
14. **Barbour EK, Nabbut NH, Hamadeh SK, Al-Nakhli HM.** 1997. Bacterial identity and characteristics in healthy and unhealthy respiratory tracts of sheep and calves. Vet Res Commun **21:**421–430. http://dx.doi.org/10.1023/A:1005855318665.
15. **Besser TE, Cassirer EF, Potter KA, VanderSchalie J, Fischer A, Knowles DP, Herndon DR, Rurangirwa FR, Weiser GC, Srikumaran S.** 2008. Association of *Mycoplasma ovipneumoniae* infection with population-limiting respiratory disease in free-ranging Rocky Mountain bighorn sheep (*Ovis canadensis canadensis*). J Clin Microbiol **46:**423–430. http://dx.doi.org/10.1128/JCM.01931-07.
16. **Queen C, Ward ACS, Hunter DL.** 1994. Bacteria isolated from nasal and tonsillar samples of clinically healthy Rocky Mountain bighorn and domestic sheep. J Wildl Dis **30:**1–7. http://dx.doi.org/10.7589/0090-3558-30.1.1.
17. **Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, Doherty C, McLachlan G.** 2015. Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate. PLoS One **10:**e0142097. http://dx.doi.org/10.1371/journal.pone.0142097.
18. **Dickson RP, Erb-Downward JR, Huffnagle GB.** 2013. The role of the bacterial microbiome in lung disease. Expert Rev Respir Med **7:**245–257. http://dx.doi.org/10.1586/ers.13.24.
19. **Davies DH, Dungworth DL, Humphreys S, Johnson AJ.** 1977. Concurrent infection of lambs with parainfluenza virus type 3 and *Pasteurella haemolytica*. N Z Vet J **25:**263–265. http://dx.doi.org/10.1080/00480169.1977.34425.
20. **Catt DM, Chengappa MM, Kadel WL, Herren CE.** 1985. Preliminary studies with a live streptomycin-dependent *Pasteurella multocida* and *Pasteurella haemolytica* vaccine for the prevention of bovine pneumonic pasteurellosis. Can J Comp Med **49:**366–371.
21. **Stockdale PHG, Langford EV, Darcel CLQ.** 1979. Experimental bovine pneumonic pasteurellosis: 1—prevention of the disease. Can J Comp Med **43:**262–271.
22. **Brogden KA, Lehmkuhl HD, Cutlip RC.** 1998. *Pasteurella haemolytica* complicated respiratory infections in sheep and goats. Vet Res **29:**233–254.
23. **Dassanayake RP, Call DR, Sawant AA, Casavant NC, Weiser GC, Knowles DP, Srikumaran S.** 2010. *Bibersteinia trehalosi* inhibits the growth of *Mannheimia haemolytica* by a proximity-dependent mechanism. Appl Environ Microbiol **76:**1008–1013. http://dx.doi.org/10.1128/AEM.02086-09.
24. **Dassanayake RP, Shanthalingam S, Herndon CN, Subramaniam R, Lawrence PK, Bavananthasivam J, Cassirer EF, Haldorson GJ, Foreyt WJ, Rurangirwa FR, Knowles DP, Besser TE, Srikumaran S.** 2010. *Mycoplasma ovipneumoniae* can predispose bighorn sheep to fatal *Mannheimia haemolytica* pneumonia. Vet Microbiol **145:**354–359. http://dx.doi.org/10.1016/j.vetmic.2010.04.011.

25. **Porter JF, Connor NK, Krueger N, Hodgson JC, Donachie W.** 1995. Predisposition of specific pathogen-free lambs to *Pasteurella haemolytica* pneumonia by *Bordetella parapertussis* infection. J Comp Pathol **112:** 381–389. http://dx.doi.org/10.1016/S0021-9975(05)80019-2.
26. **Scheerlinck JPY, Snibson KJ, Bowles VM, Sutton P.** 2008. Biomedical applications of sheep models: from asthma to vaccines. Trends Biotechnol **26:**259–266. http://dx.doi.org/10.1016/j.tibtech.2008.02.002.
27. **Meeusen EN, Snibson KJ, Hirst SJ, Bischof RJ.** 2009. Sheep as a model species for the study and treatment of human asthma and other respiratory diseases. Drug Discov Today Dis Models **6:**101–106. http://dx.doi.org/10.1016/j.ddmod.2009.12.002.
28. **Collie D, Govan J, Wright S, Thornton E, Tennant P, Smith S, Doherty C, McLachlan G.** 2013. A lung segmental model of chronic *Pseudomonas* infection in sheep. PLoS One **8:**e67677. http://dx.doi.org/10.1371/journal.pone.0067677.
29. **Berry D, Ben Mahfoudh K, Wagner M, Loy A.** 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl Environ Microbiol **77:**7846–7849. http://dx.doi.org/10.1128/AEM.05220-11.
30. **Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD.** 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Appl Environ Microbiol **79:**5112–5120. http://dx.doi.org/10.1128/AEM.01043-13.
31. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J **17:**10–12. http://dx.doi.org/10.14806/ej.17.1.200.
32. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:**7537–7541. http://dx.doi.org/10.1128/AEM.01541-09.
33. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics **27:** 2194–2200. http://dx.doi.org/10.1093/bioinformatics/btr381.
34. **Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO.** 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res **41:** D590–D596. http://dx.doi.org/10.1093/nar/gks1219.
35. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72:**5069–5072. http://dx.doi.org/10.1128/AEM.03006-05.
36. **Werner JJ, Koren O, Hugenholtz P, DeSantis TZ, Walters WA, Caporaso JG, Angenent LT, Knight R, Ley RE.** 2012. Impact of training sets on classification of high-throughput bacterial 16S rRNA gene surveys. ISME J **6:**94–103. http://dx.doi.org/10.1038/ismej.2011.82.
37. **Yue JC, Clayton MK.** 2005. A similarity measure based on species proportions. Commun Stat Theory Methods **34:**2123–2131. http://dx.doi.org/10.1080/STA-200066418.
38. **Excoffier L, Smouse PE, Quattro JM.** 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics **131:** 479–491.
39. **Paulson J, Pop M, Bravo H.** 2011. Metastats: an improved statistical method for analysis of metagenomic data. Genome Biol **12**(Suppl 1)**:**P17. http://dx.doi.org/10.1186/gb-2011-12-s1-p17.
40. **Esty WW.** 1986. The efficiency of Good's nonparametric coverage estimator. Ann Stat **14:**1257–1260. http://dx.doi.org/10.1214/aos/1176350066.
41. **Dufrêne M, Legendre P.** 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr **67:**345–366. http://dx.doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2.
42. **Efron B, Tibshirani RJ.** 1993. An introduction to the bootstrap: monographs on statistics and applied probability, vol 57. Taylor and Francis, Abingdon, United Kingdom.
43. **Dixon P.** 2003. Vegan, a package of R functions for community ecology. J Veg Sci **14:**927–930. http://dx.doi.org/10.1111/j.1654-1103.2003.tb02228.x.
44. **Anderson M, Gorley RN, Clarke RK.** 2008. Permanova+ for primer: guide to software and statistical methods. Primer-E, Plymouth, United Kingdom.
45. **Schwaiger K, Hölzel C, Mayer M, Bauer J.** 2010. Notes on the almost unknown genus *Jeotgalicoccus*. Lett Appl Microbiol **50:**441–444. http://dx.doi.org/10.1111/j.1472-765X.2010.02811.x.
46. **Malmuthuge N, Griebel PJ, Guan LL.** 2014. Taxonomic identification of commensal bacteria associated with the mucosa and digesta throughout the gastrointestinal tracts of preweaned calves. Appl Environ Microbiol **80:**2021–2028. http://dx.doi.org/10.1128/AEM.03864-13.
47. **Dewhirst FE, Klein EA, Thompson EC, Blanton JM, Chen T, Milella L, Buckley CMF, Davis IJ, Bennett M-L, Marshall-Jones ZV.** 2012. The canine oral microbiome. PLoS One **7:**e36067. http://dx.doi.org/10.1371/journal.pone.0036067.
48. **Fallschissel K, Klug K, Kämpfer P, Jäckel U.** 2010. Detection of airborne bacteria in a German turkey house by cultivation-based and molecular methods. Ann Occup Hyg **54:**934–943. http://dx.doi.org/10.1093/annhyg/meq054.
49. **Martin E, Fallschissel K, Kämpfer P, Jäckel U.** 2010. Detection of *Jeotgalicoccus* spp. in poultry house air. Syst Appl Microbiol **33:**188–192. http://dx.doi.org/10.1016/j.syapm.2010.03.008.
50. **Braem G, De Vliegher S, Verbist B, Piessens V, Van Coillie E, De Vuyst L, Leroy F.** 2013. Unraveling the microbiota of teat apices of clinically healthy lactating dairy cows, with special emphasis on coagulase-negative staphylococci. J Dairy Sci **96:**1499–1510. http://dx.doi.org/10.3168/jds.2012-5493.
51. **Osés SM, Diez AM, Melero B, Luning PA, Jaime I, Rovira J.** 2013. Characterization by culture-dependent and culture-independent methods of the bacterial population of suckling-lamb packaged in different atmospheres. Food Microbiol **36:**216–222. http://dx.doi.org/10.1016/j.fm.2013.05.005.
52. **Leng J, Cheng YM, Zhang CY, Zhu RJ, Yang SL, Gou X, Deng WD, Mao HM.** 2012. Molecular diversity of bacteria in Yunnan yellow cattle (*Bos taurs*) from Nujiang region, China. Mol Biol Rep **39:**1181–1192. http://dx.doi.org/10.1007/s11033-011-0848-5.
53. **Dungan RS.** 2012. Use of a culture-independent approach to characterize aerosolized bacteria near an open-freestall dairy operation. Environ Int **41:**8–14. http://dx.doi.org/10.1016/j.envint.2011.12.004.
54. **Lawrence PK, Shanthalingam S, Dassanayake RP, Subramaniam R, Herndon CN, Knowles DP, Rurangirwa FR, Foreyt WJ, Wayman G, Marciel AM, Highlander SK, Srikumaran S.** 2010. Transmission of *Mannheimia haemolytica* from domestic sheep (*Ovis aries*) to bighorn sheep (*Ovis canadensis*): unequivocal demonstration with green fluorescent protein-tagged organisms. J Wildl Dis **46:**706–717. http://dx.doi.org/10.7589/0090-3558-46.3.706.
55. **Blackall PJ, Bojesen AM, Christensen H, Bisgaard M.** 2007. Reclassification of *Pasteurella trehalosi* as *Bibersteinia trehalosi* gen. nov., comb. nov. Int J Syst Evol Microbiol **57:**666–674. http://dx.doi.org/10.1099/ijs.0.64521-0.
56. **Nørskov-Lauritsen N, Kilian M.** 2006. Reclassification of *Actinobacillus actinomycetemcomitans*, *Haemophilus aphrophilus*, *Haemophilus paraphrophilus* and *Haemophilus segnis* as *Aggregatibacter actinomycetemcomitans* gen. nov., comb. nov., *Aggregatibacter aphrophilus* comb. nov. and *Aggregatibacter segnis* comb. nov., and emended description of *Aggregatibacter aphrophilus* to include V factor-dependent and V factor-independent isolates. Int J Syst Evol Microbiol **56:**2135–2146. http://dx.doi.org/10.1099/ijs.0.64207-0.
57. **Smith CB, Golden CA, Kanner RE, Renzetti AD.** 1976. *Haemophilus influenzae* and *Haemophilus parainfluenzae* in chronic obstructive pulmonary disease. Lancet **i:**1253–1255.
58. **Kittelmann S, Kirk MR, Jonker A, McCulloch A, Janssen PH.** 2015. Buccal swabbing as a noninvasive method to determine bacterial, archaeal, and eukaryotic microbial community structures in the rumen. Appl Environ Microbiol **81:**7470–7483. http://dx.doi.org/10.1128/AEM.02385-15.
59. **Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L, Pachter L, Moffatt MF, Cookson WOC.** 2010. Disordered microbial communities in asthmatic airways. PLoS One **5:**e8578. http://dx.doi.org/10.1371/journal.pone.0008578.
60. **Bouljihad M, Leipold HW.** 1994. An ultrastructural study of pulmonary bronchiolar and alveolar epithelium in sheep. Zentralbl Veterinarmed A **41:**573–586. http://dx.doi.org/10.1111/j.1439-0442.1994.tb00125.x.
61. **Griebel PJ, Entrican G, Rocchi M, Beskorwayne T, Davis WC.** 2007. Cross-reactivity of mAbs to human CD antigens with sheep leukocytes.

Vet Immunol Immunopathol **119:**115–122. http://dx.doi.org/10.1016/j.vetimm.2007.06.015.
62. **Abraham W, Gill A, Ahmed A, Sielczak M, Lauredo I, Botinnikova Y, Lin K-C, Pepinsky B, Leone D, Lobb R, Adams S.** 2000. A small-molecule, tight-binding Inhibitor of the integrin $\alpha_4\beta_1$ blocks antigen-induced airway responses and inflammation in experimental asthma in sheep. Am J Respir Crit Care Med **162:**603–611. http://dx.doi.org/10.1164/ajrccm.162.2.9911061.
63. **Abraham WM, Sielczak MW, Wanner A, Perruchoud AP, Blinder L, Stevenson JS, Ahmed A, Yerger LD.** 1988. Cellular markers of inflammation in the airways of allergic sheep with and without allergen-induced late responses. Am Rev Respir Dis **138:**1565–1571. http://dx.doi.org/10.1164/ajrccm/138.6.1565.
64. **Snibson KJ, Bischof RJ, Slocombe RF, Meeusen EN.** 2005. Airway remodelling and inflammation in sheep lungs after chronic airway challenge with house dust mite. Clin Exp Allergy **35:**146–152. http://dx.doi.org/10.1111/j.1365-2222.2005.02137.x.
65. **Liravi B, Piedrafita D, Nguyen G, Bischof R.** 2015. Dynamics of IL-4 and IL-13 expression in the airways of sheep following allergen challenge. BMC Pulm Med **15:**101. http://dx.doi.org/10.1186/s12890-015-0097-9.
66. **Esther CR, Jr, Boucher RC, Johnson MR, Ansede JH, Donn KH, O'Riordan TG, Ghio AJ, Hirsh AJ.** 2014. Airway drug pharmacokinetics via analysis of exhaled breath condensate. Pulm Pharmacol Ther **27:**76–82. http://dx.doi.org/10.1016/j.pupt.2013.07.009.
67. **Alton EWFW, Baker A, Baker E, Boyd AC, Cheng SH, Coles RL, Collie DDS, Davidson H, Davies JC, Gill DR, Gordon C, Griesenbach U, Higgins T, Hyde SC, Innes JA, McCormick D, McGovern M, McLachlan G, Porteous DJ, Pringle I, Scheule RK, Shaw DJ, Smith S, Sumner-Jones SG, Tennant P, Vrettou C.** 2013. The safety profile of a cationic lipid-mediated cystic fibrosis gene transfer agent following repeated monthly aerosol administration to sheep. Biomaterials **34:**10267–10277. http://dx.doi.org/10.1016/j.biomaterials.2013.09.023.
68. **El Mays TY, Choudhury P, Leigh R, Koumoundouros E, Van der Velden J, Shrestha G, Pieron CA, Dennis JH, Green FHY, Snibson KJ.** 2014. Nebulized perflubron and carbon dioxide rapidly dilate constricted airways in an ovine model of allergic asthma. Respir Res **15:**98. http://dx.doi.org/10.1186/s12931-014-0098-x.
69. **Popp JD, McAllister TA, Kastelic JP, Majak W, Ayroud M, VanderKop MA, Karren D, Yost GS, Cheng KJ.** 1998. Effect of melengestrol acetate on development of 3-methylindole-induced pulmonary edema and emphysema in sheep. Can J Vet Res **62:**268–274.
70. **Tsai L, Hoffman A, Berger R, Ingenito E.** 2004. Bronchoscopic lung volume reduction in a sheep model of heterogeneous emphysema. J Bronchol Interv Pulmonol **11:**83–86.
71. **Ingenito EP, Tsai L, Murthy S, Tyagi S, Mazan M, Hoffman A.** 2012. Autologous lung-derived mesenchymal stem cell transplantation in experimental emphysema. Cell Transplant **21:**175–189. http://dx.doi.org/10.3727/096368910X550233.
72. **Pohlmann JR, Akay B, Camboni D, Koch KL, Mervak BM, Cook KE.** 2012. A low-mortality model of chronic pulmonary hypertension in sheep. J Surg Res **175:**44–48. http://dx.doi.org/10.1016/j.jss.2011.02.049.
73. **McDonald JW, Ali M, Morgan E, Townsend ER, Cooper JD.** 1983. Thromboxane synthesis by sources other than platelets in association with complement-induced pulmonary leukostasis and pulmonary hypertension in sheep. Circ Res **52:**1–6. http://dx.doi.org/10.1161/01.RES.52.1.1.
74. **Weimann J, Ullrich R, Hromi J, Fujino Y, Clark MWH, Bloch KD, Zapol WM.** 2000. Sildenafil is a pulmonary vasodilator in awake lambs with acute pulmonary hypertension. Anesthesiology **92:**1702–1712. http://dx.doi.org/10.1097/00000542-200006000-00030.
75. **Julien M, Hoeffel JM, Flick MR.** 1986. Oleic-acid lung injury in sheep. J Appl Physiol **60:**433–440.
76. **Enkhbaatar P, Murakami K, Shimoda K, Mizutani A, Traber L, Phillips GB, Parkinson JF, Cox R, Hawkins H, Herndon D, Traber D.** 2003. The inducible nitric oxide synthase inhibitor BBS-2 prevents acute lung injury in sheep after burn and smoke inhalation injury. Am J Respir Crit Care Med **167:**1021–1026. http://dx.doi.org/10.1164/rccm.200209-1031PP.
77. **Tsuno K, Prato P, Kolobow T.** 1990. Acute lung injury from mechanical ventilation at moderately high airway pressures. J Appl Physiol **69:**956–961.
78. **Gibbons MM, Dang X, Adkins M, Powell B, Chan P.** 2015. Finite element modeling of blast lung injury in sheep. J Biomech Eng **137:**041002. http://dx.doi.org/10.1115/1.4029181.
79. **Olivier A, Gallup J, De Macedo MMMA, Varga SM, Ackermann M.** 2009. Human respiratory syncytial virus A2 strain replicates and induces innate immune responses by respiratory epithelia of neonatal lambs. Int J Exp Pathol **90:**431–438. http://dx.doi.org/10.1111/j.1365-2613.2009.00643.x.
80. **Morris DL, Clarkson MJ, Stallbaumer MF, Pritchard J, Jones RS, Chinnery JB.** 1985. Albendazole treatment of pulmonary hydatid cysts in naturally infected sheep: a study with relevance to the treatment of hydatid cysts in man. Thorax **40:**453–458. http://dx.doi.org/10.1136/thx.40.6.453.
81. **Enkhbaatar P, Joncam C, Traber L, Nakano Y, Wang J, Lange M, Connelly R, Kulp G, Saunders F, Huda R, Cox R, Schmalstieg F, Herndon D, Traber D.** 2008. Novel ovine model of methicillin-resistant *Staphylococcus aureus*-induced pneumonia and sepsis. Shock **29:**642–649.
82. **Warner AE, Decamp MM, Molina RM, Brain JD.** 1988. Pulmonary removal of circulating endotoxin results in acute lung injury in sheep. Lab Invest **59:**219–230.
83. **Kramer BW, Kramer S, Ikegami M, Jobe AH.** 2002. Injury, inflammation, and remodeling in fetal sheep lung after intra-amniotic endotoxin. Am J Physiol Lung Cell Mol Physiol **283:**L452–L459. http://dx.doi.org/10.1152/ajplung.00407.2001.
84. **Engel M, Nowacki RE, Boden P, Reiss L, Uhlig S, Reynaert N, Gopal P, Wouters EM, Willems CMP, Kloosterboer N, Wolfs TAM, Zimmermann LI, Vos G, Kramer B.** 2015. The effects of dexamethasone and oxygen in ventilated adult sheep with early phase acute respiratory distress syndrome. Lung **193:**97–103. http://dx.doi.org/10.1007/s00408-014-9670-x.
85. **Langer T, Vecchi V, Belenkiy SM, Cannon JW, Chung KK, Cancio LC, Gattinoni L, Batchinsky AI.** 2014. Extracorporeal gas exchange and spontaneous breathing for the treatment of acute respiratory distress syndrome: an alternative to mechanical ventilation? Crit Care Med **42:**e211–e220. http://dx.doi.org/10.1097/CCM.0000000000000121.
86. **Bégin R, Rola-Pleszczynski M, Massé S, Lemaire I, Sirois P, Boctor M, Nadeau D, Drapeau G, Bureau MA.** 1983. Asbestos-induced lung injury in the sheep model: the initial alveolitis. Environ Res **30:**195–210. http://dx.doi.org/10.1016/0013-9351(83)90180-9.
87. **Lee TC, Gold LI, Reibman J, Aston C, Bégin R, Rom WN, Jagirdar J.** 1997. Immunohistochemical localization of transforming growth factor-β and insulin-like growth factor-I in asbestosis in the sheep model. Int Arch Occup Environ Health **69:**157–164. http://dx.doi.org/10.1007/s004200050132.
88. **Sebastien P, Begin R, Masse S.** 1990. Mass, number and size of lung fibres in the pathogenesis of asbestosis in sheep. J Exp Pathol **71:**1–10.
89. **Palmarini M, Fan H, Sharp JM.** 1997. Sheep pulmonary adenomatosis: a unique model of retrovirus associated lung cancer. Trends Microbiol **5:**478–483. http://dx.doi.org/10.1016/S0966-842X(97)01162-1.
90. **Leroux C, Girard N, Cottin V, Greenland T, Mornex JF, Archer F.** 2007. Jaagsiekte sheep retrovirus (JSRV): from virus to lung cancer in sheep. Vet Res **38:**211–228. http://dx.doi.org/10.1051/vetres:2006060.
91. **Jorth P, Staudinger BJ, Wu X, Hisert KB, Hayden H, Garudathri J, Harding CL, Radey MC, Rezayat A, Bautista G, Berrington WR, Goddard AF, Zheng C, Angermeyer A, Brittnacher MJ, Kitzman J, Shendure J, Fligner CL, Mittler J, Aitken ML, Manoil C, Bruce JE, Yahr TL, Singh PK.** 2015. Regional isolation drives bacterial diversification within cystic fibrosis lungs. Cell Host Microbe **18:**307–319. http://dx.doi.org/10.1016/j.chom.2015.07.006.
92. **Ingenito EP, Solway J, McFadden ER, Pichurko B, Bowman HF, Michaels D, Drazen JM.** 1987. Indirect assessment of mucosal surface temperatures in the airways: theory and tests. J Appl Physiol **63:**2075–2083.
93. **Martin CJ, Cline F, Marshall H.** 1953. Lobar alveolar gas concentrations: effect of body position. J Clin Invest **32:**617–621. http://dx.doi.org/10.1172/JCI102772.
94. **West JB.** 1978. Regional differences in the lung. Chest **74:**426–437. http://dx.doi.org/10.1378/chest.74.4.426.
95. **Robertson HT, Hlastala MP.** 2007. Microsphere maps of regional blood flow and regional ventilation. J Appl Physiol **102:**1265–1272.
96. **Glenny RW.** 2009. Determinants of regional ventilation and blood flow in the lung. Intensive Care Med **35:**1833–1842. http://dx.doi.org/10.1007/s00134-009-1649-3.

97. **Marsland BJ, Gollwitzer ES.** 2014. Host-microorganism interactions in lung diseases. Nat Rev Immunol **14:**827–835. http://dx.doi.org/10.1038/nri3769.
98. **Duncan SH, Louis P, Thomson JM, Flint HJ.** 2009. The role of pH in determining the species composition of the human colonic microbiota. Environ Microbiol **11:**2112–2122. http://dx.doi.org/10.1111/j.1462-2920.2009.01931.x.
99. **Roth RR, James WD.** 1988. Microbial ecology of the skin. Annu Rev Microbiol **42:**441–464. http://dx.doi.org/10.1146/annurev.mi.42.100188.002301.
100. **Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW.** 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol **12:**87. http://dx.doi.org/10.1186/s12915-014-0087-z.
101. **Jones FS.** 1922. The source of the microorganisms in the lungs of normal animals. J Exp Med **36:**317–328. http://dx.doi.org/10.1084/jem.36.3.317.
102. **Ward DV, Gevers D, Giannoukos G, Earl AM, Methe BA, Sodergren E, Feldgarden M, Ciulla DM, Tabbaa D, Arze C, Appelbaum E, Aird L, Anderson S, Ayvaz T, Belter E, Bihan M, Bloom T, Crabtree J, Courtney L, Carmichael L, Dooling D, Erlich RL, Farmer C, Fulton L, Fulton R, Gao HY, Gill JA, Haas BJ, Hemphill L, Hall O, Hamilton SG, Hepburn TA, Lennon NJ, Joshi V, Kells C, Kovar CL, Kalra D, Li K, Lewis L, Leonard S, Muzny DM, Mardis E, Mihindukulasuriya K, Magrini V, O'Laughlin M, Pohl C, Qin X, Ross K, Ross MC, Rogers YHA, et al.** 2012. Evaluation of 16S rRNA gene-based community profiling for human microbiome research. PLoS One **7:**e39315. http://dx.doi.org/10.1371/journal.pone.0039315.
103. **Schloss PD, Gevers D, Westcott SL.** 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One **6:**e27310. http://dx.doi.org/10.1371/journal.pone.0027310.
104. **Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J.** 2014. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. PLoS One **9:**e94249. http://dx.doi.org/10.1371/journal.pone.0094249.
105. **Pinto AJ, Raskin L.** 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing data sets. PLoS One **7:**e43093. http://dx.doi.org/10.1371/journal.pone.0043093.
106. **Kembel SW, Wu M, Eisen JA, Green JL.** 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol **8:**e1002743. http://dx.doi.org/10.1371/journal.pcbi.1002743.

3.3 Discussion

The spatial variation observed in the sheep lung microbiota, even across very small distances, suggests that it is unlikely to be possible to characterise the entire lung microbiota of an individual sheep using protected specimen brushings. For some studies it may therefore be desirable to perform lung washes which would be able to sample a greater area and give a better representation of the lung. However, performing whole lung washes while minimising contamination in live animals would be difficult if not impossible. Additionally, if sampling to investigate the effect of certain treatments or diseases which may have localised lung effects, protected specimen brushings may still prove to be more useful.

The predominant types of bacteria identified in the lungs of sheep during this study were different to those previously reported for humans. However, we were limited by the number of sheep used and it is possible that these six sheep, who were all pen mates, were not representative of other sheep populations or breeds. The high level of inter-individual variation observed also merits a study involving a larger number of animals.

Chapter 4: Comparing Microbiotas in the Upper Aerodigestive and Lower Respiratory Tracts of Lambs

4.1 Introduction and aims

The results in Chapter 3 give a first glimpse into the types of bacteria which may inhabit the sheep lung. We identified differences between the microbiota we observed in sheep and that which had previously been described in humans. This led us to question whether, as in healthy humans (23, 39), the upper and lower respiratory tract microbiotas are similar to one another or whether there are differences between these niches which may point to the sheep lung microbiota being formed through different mechanisms than are responsible in humans.

The following manuscript was submitted to Microbiome on 19/07/2016. The inclusion of this submitted manuscript as part of this thesis does not conflict with any Biomed Central editorial policies**.** The manuscript characterises the composition of the sheep lung microbiota in a larger number of animals than was used in Chapter 3. Upper and lower respiratory tract samples are also compared in order to explore the differences between the microbiota at these niches. In this study, all work was performed by myself except as further specified: lung fluid samples were collected by myself with the help of Steven Wright and David Collie. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording.

4.2 Submitted manuscript: Comparing microbiotas in the upper aerodigestive and lower respiratory tracts of lambs

Laura Glendinning[1] (laura.glendinning@roslin.ed.ac.uk), David Collie[1] (david.collie@roslin.ed.ac.uk), Steven Wright[1] (steven.wright@roslin.ed.ac.uk), Kenny MD Rutherford[2] (Kenny.Rutherford@sruc.ac.uk), Gerry McLachlan[1] (gerry.mclachlan@roslin.ed.ac.uk)

This submitted manuscript is included here as a monograph thesis as described in 'GUIDANCE FOR THE INCLUSION OF PUBLICATIONS IN POSTGRADUATE RESEARCH THESES' -1.6. This manuscript was submitted to Microbiome on 19/07/2016.

1 The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, Midlothian, United Kingdom, EH25 9RG

2Animal Behaviour and Welfare, Animal and Veterinary Sciences Research Group, SRUC, West Mains Rd., Edinburgh, Midlothian, United Kingdom, EH9 3JG

4.2.1 Abstract

Background: Recently, the importance of the lung microbiota during health and disease has been examined in humans and in small animal models. Whilst sheep have been proposed as an appropriate large animal model for studying the pathophysiology of a number of important human respiratory diseases, it is clearly important to continually define the limits of agreement between these systems as new concepts emerge. In humans it has recently been established that the lung microbiota is seeded by microbes from the oral cavity. We sought to determine whether the same was true in sheep.

Results: We took lung fluid and upper aerodigestive tract (oropharyngeal) swab samples from forty lambs (seven weeks old). DNA extraction was performed and the V2-V3 region of the 16S rRNA gene was amplified by PCR then sequenced via Illumina MiSeq. Oropharyngeal swabs were either dominated by bacteria commonly associated with the rumen or by bacteria commonly associated with the upper aerodigestive tract. Lung microbiota samples did not resemble either upper aerodigestive tract samples or reagent only controls. Some rumen associated bacteria were found in lung fluids, indicating that inhalation of ruminal bacteria does occur. We also identified several bacteria which were significantly more abundant in lung fluids than in upper aerodigestive tract swabs, the most predominant of which was classified as *Staphylococcus equorum*.

Conclusions: In contrast to humans, we found that that the lung microbiota of lambs is dissimilar to that of the upper aerodigestive tract and we suggest that this may be related to physiological differences between sheep and humans. Understanding the comparative physiology underlying differences in lung microbiota between species will provide a foundation upon which to interpret changes associated with disease and/or environment.

Keywords: lung, microbiota, sheep, lambs, oropharynx, rumen, 16S

4.2.2 Background

The use of 16S ribosomal RNA (rRNA) gene sequencing has facilitated the study of difficult to culture, low biomass microbial communities present in the lower respiratory tract. The impact of the lung microbiota on human health is a rapidly growing area of research. In order to understand this impact it is important to also understand the lung microbiota dynamics during health and to include healthy controls in disease studies. To achieve this, the majority of previous studies have relied on human volunteers.

However, many individuals are hesitant to participate in research bronchoscopy due to the percieved inconvenience and a fear of complications (210), despite the low risk involved. Mice and rats have been used to explore the relationship between the lung microbiota and airway inflammation (211-213), microbiota at different body sites (214), the environment (51) acute lung injury (98) and antibiotic (215) and corticosteroid exposure (216). However, rodents are of limited use when exploring spatial or longitudinal lung microbiota dynamics due to their small lung size. Recognising the utility of large animal models in this regard, and the anatomical and immunological relevance of sheep as models (11, 12, 217, 218) our group has previously used this species to explore the changes in the lung microbiota upon *Pseudomonas aeruginosa* infection (115) and to explore the spatial variability present within the healthy lung (1).

Subclinical microaspiration of pharyngeal secretions is a feature of health and this can contribute to the lung microbiota composition (14) and the microbiome of the human lungs more closely resembles that of the mouth than the nose or the lower gastrointestinal tract (39). It is not yet known whether the same relationship holds for species other than humans. In ruminating sheep, where the oropharynx is exposed to ruminal contents on a frequent basis one would anticipate that lung microbiota would similarly reflect this influence. In this paper we find that the presence of rumen-like bacteria in the upper aerodigestive tract is correlated with changes in the lung microbiota and rumen type bacteria are present in lamb lungs. We also identify bacteria which are more indicative of the lungs than the oropharynx, indicating that the presence of the sheep lung microbiota is not merely due to passive diffusion of microbes from the upper aerodigestive tract.

4.2.3 Methods

Animals and sampling

Scottish Mule X Suffolk lambs (20 males and 20 females), raised on pasture from 48 hours after birth, were used in this study. These lambs were part of a study on the animal welfare implications of prenatal stress which was approved by Scotland's Rural College's (SRUC) Animal Experiments Committee and was conducted under Home Office licence. At seven weeks old (mean age = 48.8 days ± 0.8 standard deviation (SD); mean weight ± SD = 20.6 kg ± 2.6 kg), lambs were euthanized by barbiturate overdose then the cadavers were transported from the farm to the dissecting room (~5 mins). Oropharyngeal swabs were taken using cotton tipped swabs (Swab Plain Wood Cotton Tip Sterile (710-0181), Copan, Italy). To prevent oral contamination, swabs were stored in protective plastic sheaths from which the swab could be advanced and retracted once it was positioned at the sampling site. Swabs were then transferred into a new plastic sheath and stored on ice.

The ventral aspect of the neck was shaved and a sterile scalpel used to incise through the skin and subcutaneous tissues to expose the ventral surface of the trachea. A sampling site was identified on the exposed ventral surface and the trachea cranial to this site was completely closed off by both string ligature and clamp placement. The selected sampling site was then heat seared and 50 ml of sterile phosphate buffered saline (PBS) was injected through the seared section into the tracheal lumen. The head and neck were oriented such that the PBS would flow caudally down the thorax. A second clamp was immediately placed caudal to the site of injection to prevent backflow, leakage and potential contamination. The lamb cadavers were then tipped so that the PBS would run caudally into their lungs and then tipped back again so that the fluid would collect in the tracheal lumen immediately caudal to the position of the second clamp. A sampling site identified on the ventral surface of the trachea was seared and a needle and syringe were used to collect the pooled fluid. On average 4 ± 1.7 ml (mean ± SD) of lung fluid was collected per animal. Lung fluid was stored on ice until further processing. Oropharyngeal swabs were sterilely cut into 500 µl PBS. Lung fluids were centrifuged at 13,000g for 5 mins (Biofuge Fresco Heraeus, Germany**)**. The supernatant was removed and the pellets were resuspended in 500 µl PBS. Oropharyngeal swabs and lung fluids were stored at -80°C until DNA extraction.

DNA extraction, amplification, and sequencing

DNA extractions using the PowerSoil DNA isolation Kit (Mo Bio, Carlsbad, USA) and quantitative PCR (qPCR) using the 16S rRNA gene qPCR primers UniF340 (5-ACTCCTACGGGAGGCAGCAGT-3) and UniR514 (5-ATTACCGCGGCTGCTGGC-3) were performed as described previously (1). Extraction kit reagent controls, consisting of reagent only extractions, were produced for every day DNA extractions were performed. PBS controls were

created by extracting DNA from 500 µl of the PBS which had also been used to wash out the lamb lungs. A mock community control was included which has been described previously (1).

A nested PCR reaction was used to produce amplicons for sequencing; this technique was chosen to reduce PCR bias caused by barcoded primers (181). The first round of PCR amplified the V1-V4 16S hypervariable regions using the primers 28F (5-GAGTTTGATCNTGGCTCAG-3) and 805R (5-GACTACCAGGGTATCTAATC-3). The conditions were: 94°C for 2 mins followed by 20 cycles of 94°C for 1 min, 55°C for 45 s, and 72°C for 1.5 mins followed by a final extension step of 72°C for 20 mins. Clean-up was performed using the AMPure XP PCR purification system (Beckman Coulter, Brea, USA).

In a previous study, we found that PCR bias in high template samples could be reduced by diluting amplicons from the first round of PCR to a similar concentration to those of lung fluid samples (1). Therefore, in this study we used our qPCR values to calculate the dilutions needed to achieve this. The second round of PCR used the barcoded V2-V3 primers 104F (5-GGCGVACGGGTGAGTAA-3) and 519R (5-GTNTTACNGCGGCKGCTG-3). The dilutions and barcoded primers used for each sample can be found in **Additional file 4.1**. The PCR conditions were: 98°C for 30 s followed by 20 cycles of 98°C for 10 s, 67°C for 30 s, and 72°C for 10 s followed by a final extension step of 72°C for 2 mins. The amplicons were again purified using the AMPure XP PCR purification system.

Bioinformatic and statistical analysis

Samples were sequenced via either Illumina MiSeq or HiSeq runs (Illumina, San Diego, USA) (**Additional file 4.1**) producing 250 base pair paired-end reads. Cutadapt (184) was used to remove primers. Quality control and taxonomic assignment of sequences was carried out within mothur (186) following a protocol created by the mothur developers (187), adjusted to suit our dataset (1). Sequences were subsampled before statistical analysis. The sequencing error rate, principal coordinate analysis graphs (PCOA); analyses of molecular variance (AMOVA); Good's coverage analyses (219); richness (Chao 1 Index) and diversity (Inverse Simpson Index) calculations and indicator analyses (220) were all calculated within mothur. Clustering of microbial communities into metacommunities was also carried out within mothur using a probabilistic modelling technique based upon work by Holmes *et al* (201). The significance of differences between the diversity and richness of groups was calculated using either the two sample t-test (normal data) or the Mann-Whitney U test (non-normal data) within Minitab 16 for Windows (Minitab, Coventry, UK). Heatmaps were constructed in R Version 3.2.2 (204) using the Vegan (208), RColorBrewer (207), gplots (205) and heatplus (206) packages. Clustering within heatmaps was performed using the Bray-Curtis dissimilarity (209). Sequences can be accessed via the Bioproject ascension number PRJNA317719.

4.2.4 Results

Quality assurance of methodology

11,878,769 sequence reads were produced in total with an average of 138,125 ± 29,306 per sample (mean ± SD). The sequencing error rate was calculated as 0.35%. The oropharyngeal swab sample from lamb 12773 was found to have very low read numbers and was therefore discarded from statistical analyses, as was its corresponding lung fluid sample. A total of 1061 OTUs were identified (**Additional file 4.2**) which were reduced to 750 after sub-sampling. All Good's coverage values were >0.999 indicating that at least 99.9% of the bacteria present in our samples are likely to have been identified. The most abundant bacterial OTUs from extraction kit reagent only controls are listed in **Table 4.1**. The similarity of the OTUs found on the 25th and 26th March 2015 is likely due to the fact that the same lot of extraction kit was used. Upon examining our data we found that lung fluid samples clustered by when they were processed (**Appendix 3**). Samples sequenced via MiSeq and HiSeq underwent DNA extraction and PCR amplification separately. We identified two OTUs which were significantly indicative ($P < 0.05$) of samples from either the HiSeq or MiSeq run which were also present in all lung fluid samples from the run they were indicative of: OTU 4 (*Pseudomonas*) and OTU 112 (Yaniellaceae). These OTUs are likely due to contamination and were therefore removed prior to analysis.

Table 4.1: Bacterial OTUs found to be >5% abundant in extraction kit reagent controls

Date of DNA extraction	OTUs	Abundance
17th July 2014	*Aerococcus*	14%
	Dermabacteraceae	12%
	Micrococcus	10%
	Enhydrobacter	9%
	Leuconostoc	6%
	Kocuria	6%
	Actinomyces	6%
25th March 2015	*Methylobacterium komagatae*	65%
	Ruminococcaceae	11%
	Methylobacterium	6%
26th March 2015	*Methylobacterium komagatae*	67%
	Methylobacterium	6%

Lamb oropharyngeal swabs cluster into two distinct community types

Oropharyngeal swabs were taken from 40 lambs. Using the Laplace approximation it was found that swabs could be partitioned into two separate groups based upon the types of bacteria present. These appeared to correspond to either oropharyngeal-type (partition 1) or rumen-type (partition 2) bacteria (**Appendix 4**). Oropharyngeal-type communities were dominated by the OTUs Pasteurellaceae (22%), *Mannheimia* (14%), *Fusobacterium* (11%), *Bibersteinia trehalosi* (8%), Neisseriaceae (7%), *Moraxella* (6%) and *Bibersteinia* (5%). Rumen-type communities were dominated by the OTUs *Prevotella* (36%), Clostridiales (11%), Ruminococcaceae (7%), Lachnospiraceae (6%) and *Butyrivibrio* (6%).

The richness (chao: non-normal data) and diversity (Inverse Simpsons: normal data) of the partitions were compared. There was no significant difference in richness or diversity between the rumen-type partition and the oropharyngeal-type partition.

Dichotomous oropharyngeal microbiota are associated with different lung community structures

The most common OTUs found in lung fluid samples were *Staphylococcus equorum* (13%), *Staphylococcus sciuri* (6%), *Mannheimia* (5%) and *Prevotella* (5%). Using the Laplace approximation, lung fluids did not cluster into more than one group. Lung fluids were then manually partitioned into the same groups as swabs. A significant difference in bacterial community structure was found between these groups (AMOVA: $P = 0.016$) and a small number of OTUs were found to be significantly different between the two groups. *Prevotella* ($P = 0.03$) and *Sphingobium* ($P = 0.039$) were significantly indicative of lambs from which rumen-type swabs were derived whereas *Paracoccus aminovorans* ($P = 0.036$) was indicative of lambs from which oropharyngeal-type swabs were derived. **Fig. 4.1 and 4.2** contain visual representations of sample clustering.

We compared the proportions of the dominant OTUs in rumen-type swabs with their corresponding proportions in lung samples. On average, these OTUs were found in the following proportions in lung samples: *Prevotella* (5%), Clostridiales (2%), Ruminococcaceae (3%), Lachnospiraceae (1%) and *Butyrivibrio* (1%).

The presence of a lung specific microbiota

Indicator species analysis determined that several OTUs were significantly more indicative of the lungs than of oropharyngeal swabs (**Table 4.2**). It is likely that reagent contamination will have had more of an impact on the lung fluid samples than on the oropharyngeal swabs, due to their lower biomass. However, when examining the indicative OTUs, the majority of samples were not found to contain the same proportions of these OTUs as the PBS controls processed alongside them (**Fig. 4.3**).

Of the indicative OTUs, by far the most common was *Staphylococcus equorum* which constituted, on average, 13.3% of the total bacteria present in lung fluids and which was only present in low numbers in controls and oropharyngeal swabs.

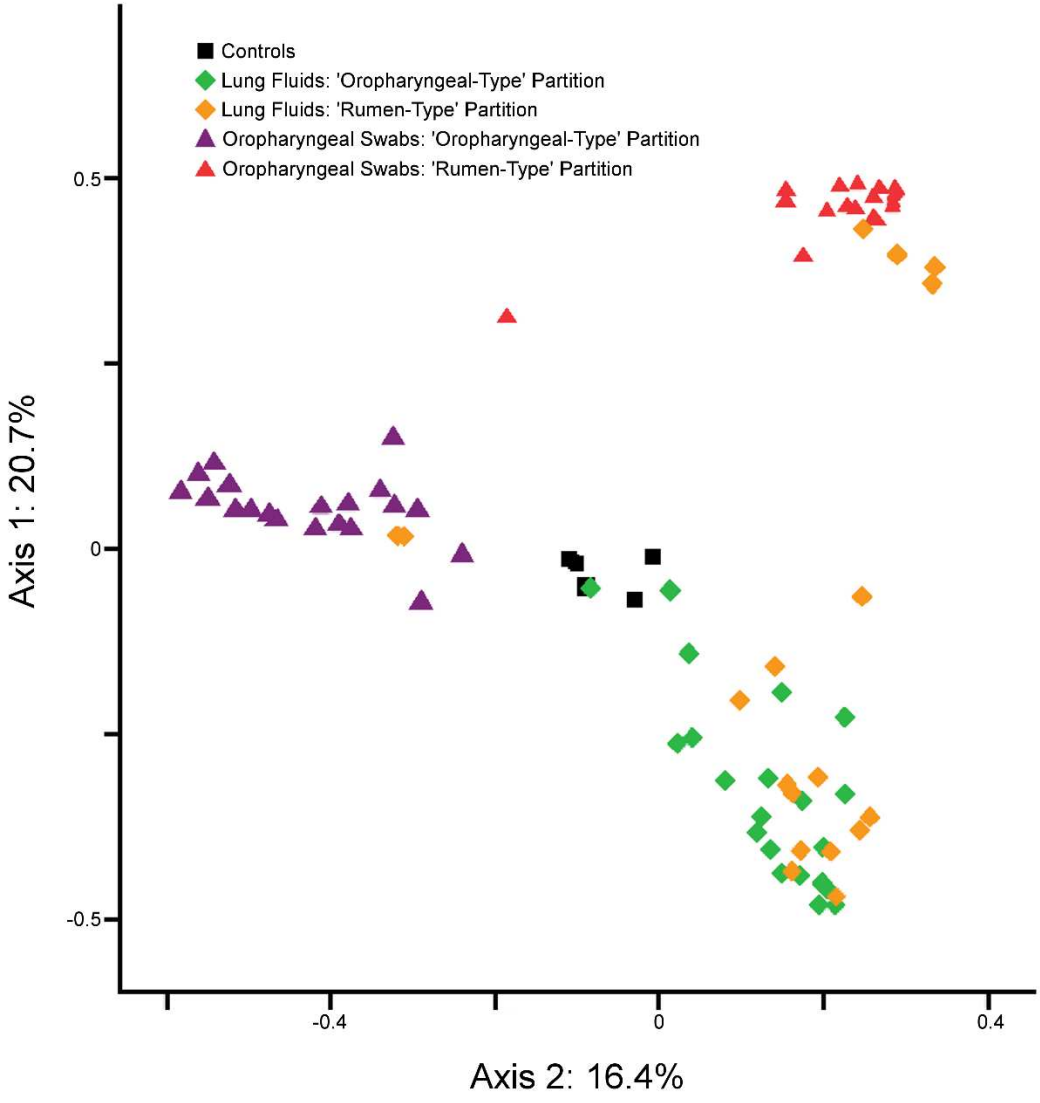


Figure 4.1: PCOA graph showing the relatedness of upper aerodigestive tract samples from lambs partitioned into 2 groups using the Laplace approximation. Lung fluids belonging to the same animals were partitioned into the same groups. Lung fluid partitions clustered significantly separately by AMOVA (P = 0.016) as did oropharyngeal swabs (P <0.001). Controls are PBS and extraction kit reagent controls.

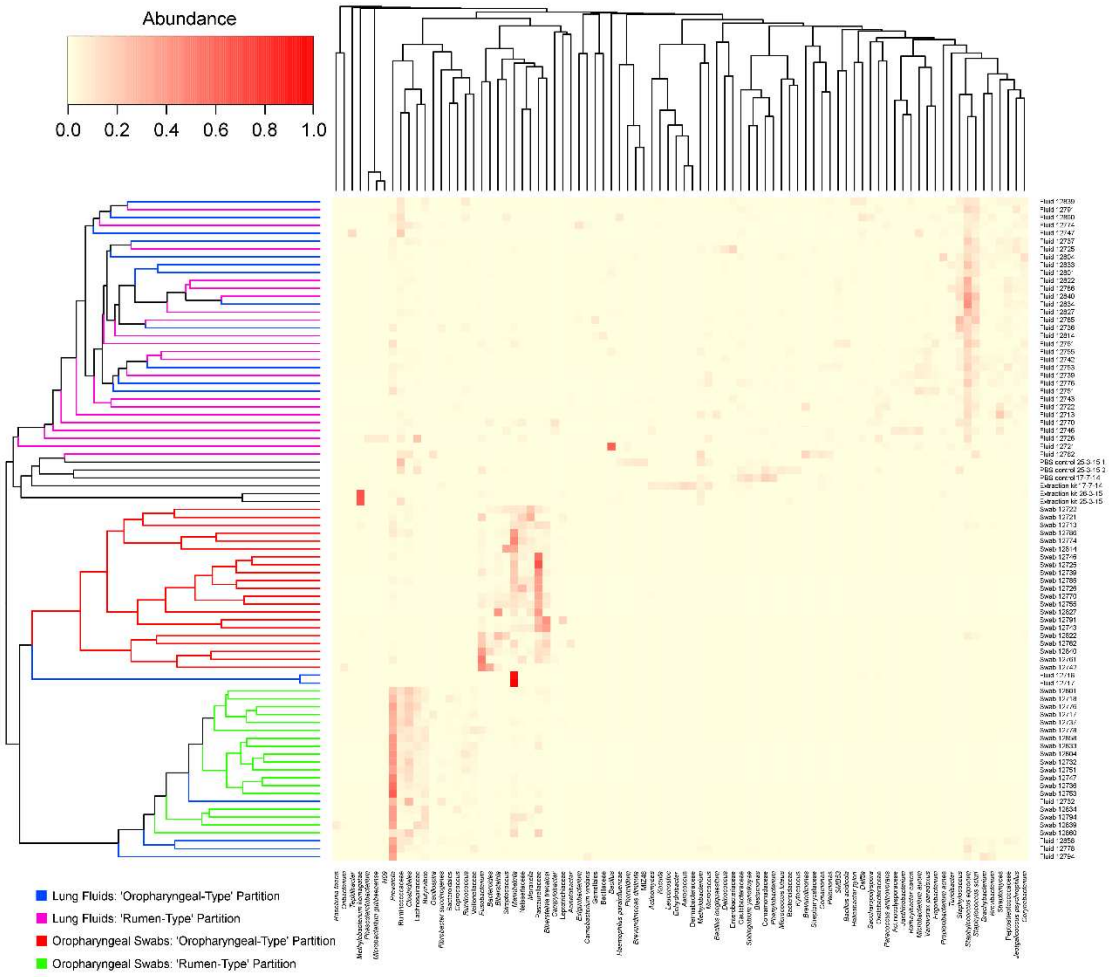


Figure 4.2: Heatmap of OTUs found in lamb lung fluids, oropharyngeal swabs, extraction kit reagent and PBS controls. OTUs were included when they were >5% abundant in at least one sample. Oropharyngeal swabs partitioned into rumen-like bacterial communities are indicated by green whereas those which were upper aerodigestive tract-like are indicated by a red line. The lung fluid samples from the oropharyngeal-like animals are indicated by blue whereas those from the rumen-type animals are indicated by pink. A larger version of this figure can be found in Appendix 5.

Table 4.2: OTUs significantly more indicative of lung fluids than oropharyngeal swabs*

Taxonomy	P-value	Average proportion in lung fluids (Mean ± SD)	Average proportion in oropharyngeal swabs (Mean ± SD)	Highest proportion in PBS controls
Corynebacterium	<0.001	1.9% ± 2.3%	0.065% ± 0.19%	0.044%
Delftia	<0.001	0.80% ± 1.7%	0% ± 0%	0%
Oxalobacteraceae	<0.001	0.96% ± 1.5%	0.012% ± 0.043%	3.0%
Pelomonas	<0.001	0.65% ± 1.1%	0.00057% ± 0.0036%	2.0%
Propionibacterium acnes	<0.001	0.84% ± 2.3%	0.0040% ± 0.020%	1.6%
Pseudomonas citronellolis	<0.001	0.51% ± 1.1%	0% ± 0%	1.0%
SMB53	<0.001	0.71% ± 1.3%	0.0045% ± 0.018%	0%
Sphingobium yanoikuyae	<0.001	0.53% ± 0.56%	0% ± 0%	13%
Staphylococcus	<0.001	3.9% ± 5.4%	0.060% ± 0.18%	8.6%
Staphylococcus equorum	<0.001	13.3% ± 9.6%	0.32% ± 0.97%	0.044%
Staphylococcus sciuri	<0.001	6.4% ± 5.3%	0.18% ± 0.59%	2.0%
Streptomyces	<0.001	2.0% ± 3.8%	0.025% ± 0.096%	0%
Brevibacterium	0.002	1.2% ± 1.4%	0.064% ± 0.24%	0%
Brachybacterium	0.006	1.0% ± 1.7%	0.035% ± 0.10%	0.022%
Peptostreptococcaceae	0.006	1.8% ± 2.2%	0.050% ± 0.11%	0.044%
Jeotgalicoccus psychrophilus	0.008	1.6% ± 2.1%	0.040% ± 0.10%	0%
Saccharopolyspora	0.009	0.52% ± 1.2%	0.0011% ± 0.0071%	0%
Janthinobacterium	0.01	0.57% ± 1.4%	0.0023% ± 0.0068%	0%
Variovorax paradoxus	0.011	1.2% ± 1.2%	0.024% ± 0.060%	0.022%
Rathayibacter caricis	0.016	0.58% ± 1.2%	0.0057% ± 0.021%	0%
Turicibacter	0.016	1.0% ± 1.8%	0.0074% ± 0.021%	0%
Micrococcus	0.017	0.77% ± 1.6%	0.0080% ± 0.029%	4.4%
Frigoribacterium	0.021	0.79% ± 1.4%	0.077% ± 0.31%	0%
Enterobacteriaceae	0.023	0.65% ± 2.6%	0.0063% ± 0.029%	2.2%
Microbacterium aurum	0.047	1.2% ± 2.8%	0.0045% ± 0.013%	0%

*OTUs which were significantly more ($P < 0.05$) indicative of lamb lung fluids than oropharyngeal swabs and which were on average >0.5% abundant in lung fluids.

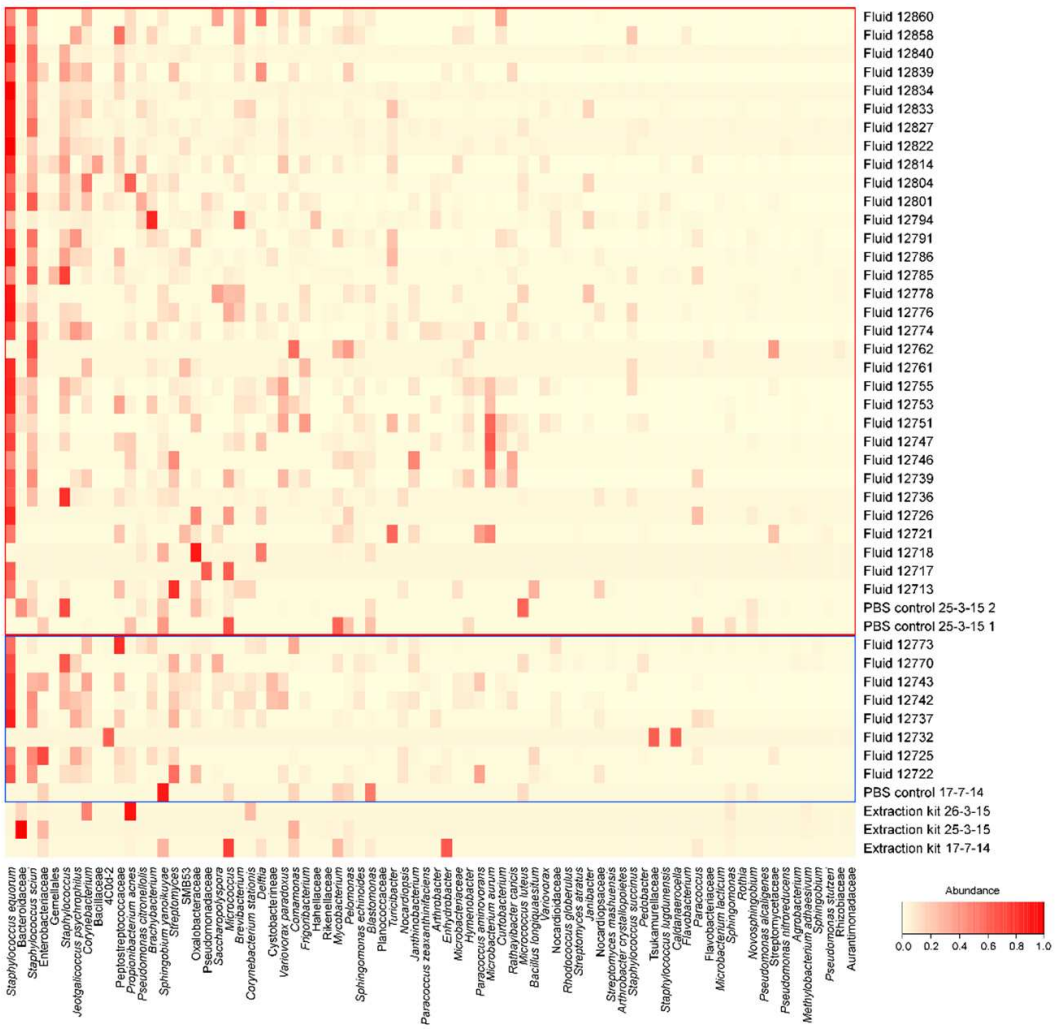


Figure 4.3: Heatmap of bacterial OTUs found to be more indicative of lamb lung fluids than oropharyngeal swabs (P <0.05). Fluid and PBS samples from which DNA was extracted on specific dates are surrounded by coloured lines: 17th July 2014 (blue) and 25th March 2015 (red). DNA extractions carried out on 26th March 2015 only comprised oropharyngeal swabs which are not included in this figure. A larger version of this figure can be found in Appendix 6.

4.2.5 Discussion

Sheep are commonly used as large animals models of the respiratory system due to the physiological and immunological similarities of their lungs to those of humans (217, 218, 221, 222). We have previously used the sheep to study both the extent of variation in the lung microbiota (1) and the direct and remote changes in the lung microbiota caused by localised *P.aeruginosa* infection and antibiotic treatment (115). As the sheep is an important agricultural animal, studies of their respiratory microbial communities may also be of interest from an animal health perspective.

It has previously been demonstrated that microaspiration of microbes from the upper aerodigestive tract is common in humans and can lead to an inflammatory phenotype (50). When microbial communities from healthy human lungs are characterised they are often found to contain microbes associated with the upper aerodigestive tract (49). The healthy human lung microbiota is thought to be formed predominantly from the neutral dispersal of these upper aerodigestive tract microbes into the lungs rather than by the differential growth of lung adapted microbial communities (23). We sought to identify whether this was also the case in sheep.

Sheep oropharyngeal swabs could be partitioned into two separate groups which were predominantly composed of OTUs identified as bacteria which are well known members of either the rumen (*Prevotella*, Clostridiales, Ruminococcaceae, Lachnospiraceae and *Butyrivibrio* (223-225)) or respiratory tract microbiotas (Pasteurellaceae, *Mannheimia*, *Fusobacterium*, *Bibersteinia trehalosi*, Neisseriaceae, *Moraxella* and *Bibersteinia* (226-228)). These bacteria were also detected in a previous study examining sheep buccal swabs (44).

It is not possible to identify whether this dichotomy reflected recent rumination, or some stochastic post-mortem leakage of rumen fluid into the oropharynx in some individuals. The lambs during this study were not weaned but were at an age when it is expected that they all would be regularly supplementing their diet with grass and would be ruminating.

Regardless of the drivers of this oropharyngeal dichotomy, the microbial communities found in the lungs were very different to those found in both the rumen- and oropharyngeal-type swabs. A large number of bacterial OTUs were found to be significantly more abundant in lung fluids in comparison to oropharyngeal swabs, including *Staphylococcus equorum* which was by far the most common bacterial OTU found. Several OTUs which are commonly associated with the rumen were also identified in lung fluids. Our lung fluid samples will have been more affected by reagent contamination than the oropharyngeal swabs due to the lower quantity of bacterial DNA present (140). However, the microbial communities found in lung fluids did not reflect the bacteria found in reagent only controls processed on the same day so the presence of bacteria in the lamb lung cannot be attributed purely to sample contamination. Nor can it be attributed to disease as no lambs showed signs of respiratory illness during the study.

There are several reasons why the microbes found in lamb lungs might not reflect those found in the upper aerodigestive tract to the same extent as is found in humans. Sheep have evolved to cope with rumination and thereby may have more efficient anatomical barriers to microaspiration (229). Physiological differences such as the horizontal disposition of the lungs, increased nasal breathing and increased saliva production (230, 231) may also contribute.

4.2.6 Conclusions

In this study we examined oropharyngeal swab and lung fluid samples taken from healthy lambs to characterise the bacterial communities present and to assess the impact of rumination on these communities. We found that oropharyngeal swabs were dominated by either rumen-type or oropharyngeal-type microbial communities. We also found that lung bacteria did not greatly resemble either rumen- or oropharyngeal-type swabs and identified several bacterial OTUs which were more indicative of lung fluids. The lungs did contain several rumen associated bacteria which may indicate that there is a certain degree of microaspiration of ruminal contents in lambs.

Sheep are not human but the opportunities that they and other large animals present, offer valuable insights into the dynamic relationship of the upper aerodigestive and lower respiratory tract microbiota in health. In the future, their value may extend to developing an understanding of the factors that predispose the upper aerodigestive tract microbiota towards disease in the lower respiratory tract.

4.2.7 Additional files

Additional file 4.1 (Additional file 4_1.xlsx): **Dataset S1**: Sample processing data for all samples.

Additional file 4.2 (Additional file 4_2.xlsx): **Dataset S2**: Full list of bacterial OTUs and taxonomies.

Appendix 3: **Figure S1:** Heatmap of OTUs found in lamb lung fluids, oropharyngeal swabs, PBS and extraction kit reagent only controls.

Appendix 4: **Table S1:** OTUs responsible for partitioning of lamb oropharyngeal swabs into two groups (using Laplace value).

Appendix 5: Larger version of **Figure 4.2**.

Appendix 6: Larger version of **Figure 4.3**.

4.3 Discussion

Based upon my findings in Chapters and 3 and 4 there are differences between the sheep and human lung microbiotas in terms of both composition and dynamics. While these differences may preclude the use of sheep as a model for some lung microbiota studies it does not preclude their use for studying the utility of different lung sampling methods. Other than when sputum is utilised, lung microbiota sampling in living individuals requires the use of bronchoscopic procedures. If a non-invasive method of lung microbiota sampling was developed this could reduce the procedural stress experienced by animals taking part in lung microbiota studies and reduce the likelihood that human volunteers would be dissuaded from participating in studies due to the perceived inconvenience and risk. Therefore, in Chapter 5 I explore the possibility of replacing invasive sampling techniques with exhaled breath condensate collection.

Chapter 5: Microbiota in Exhaled Breath Condensate and the Lung

5.1 Introduction and aims

In this Chapter I examine whether exhaled breath condensate samples (EBC) produce equivalent lung microbiota samples to those taken via protected specimen brushings, a technique which I previously used in Chapter 3. I also wanted to examine whether the healthy lung microbiota could be manipulated via antibiotic treatment and whether EBC samples were able to detect changes caused by treatment. In this study, all work was performed by myself except as further specified: colistimethate sodium (CMS) delivery and EBC collection in conscious animals were performed by myself and Gerry McLachlan. Anaesthesia was induced and monitored by Steven Wright, David Collie and Peter Tennant. EBC collection in anaesthetised animals was performed by myself, David Collie and Peter Tennant. Bronchoscopic procedures were carried out by David Collie and Peter Tennant. Blood was collected post mortem by Steven Wright, David Collie and Gerry McLachlan. HPLC was performed by Andy Gill. DNA libraries were sequenced by Edinburgh Genomics. Manuscript authors had input on the paper design and wording.

Prior to this study, in order to verify that we were able to deliver bactericidal levels of CMS to the sheep epithelial lining fluid via nebulisation we treated one sheep with 2,000,000 IU nebulised CMS once per day for seven days (as described in Section 2.7: Dosage regime 1). One month prior to treatment, BAL fluid samples were collected in order to perform differential cell counts. Cell counts were within normal ranges, indicating that it is unlikely that infectious lung disease was present (counts performed by Steven Wright). 24 hours after the final CMS dose, BAL fluid samples were taken from the RC, RCD and LCD lung segments. 48 hours after the final CMS dose, further BAL fluid samples were taken from the RA, RVD1 and LVD2 lung segments. Serum samples were also taken at 24 and 48 hours post-treatment in order to calculate the dilution factor of the epithelial lining fluid in the BAL fluid. The concentrations of colistin in BAL fluid were measured by HPLC (**Table 5.1**). The sheep did not show any adverse side-effects from treatment.

The average colistin concentration in the lungs post-treatment at 24 hours was 154.24 μg/ml and at 48 hours was 92.61 μg/ml. The lowest concentration found at either time-point was 50.76 ng/μl in the RA segment at 48 hours. This is a higher concentration than has previously been shown to successfully treat 4/5 of patients with ventilator-associated tracheobronchitis caused by polymyxin-only susceptible Gram negative bacteria (232) and is higher than the minimum inhibitory concentration values reported for susceptible Gram negative bacteria (≤2 μg/ml,(233)). This dosage regime was changed (see Section 2.7: Dosage regime 2) for the following study in order to reduce the time sheep would need to be restrained and separated from the flock, thereby reducing stress in our animals.

Table 5.1: Colistin A levels in epithelial lining fluid after CMS treatment

Lung segment	Colistin concentration (BAL fluid) (Mean ± SD)	Dilution factor	Colistin concentration (Epithelial lining fluid) (Mean)
RC (24 hours post-treatment)	2.58 ± 0.058 ng/µl	21.6	55.73 ng/µl
RCD (24 hours post-treatment)	4.37 ± 0.12 ng/µl	21.6	94.34 ng/µl
LCD (24 hours post-treatment)	9.62 ± 0.14 ng/µl	32.5	312.65 ng/µl
RA (48 hours post-treatment)	2.63 ± 0.34 ng/µl	19.3	50.76 ng/µl
RVD1 (48 hours post-treatment)	6.58 ± 0.21 ng/µl	14.5	95.41 ng/µl
LVD2 (48 hours post-treatment)	9.08 ± 0.77 ng/µl	14.5	131.66 ng/µl

5.2 Research paper

This research was published as 'Microbiota in Exhaled Breath Condensate and the Lung' in Applied and Environmental Microbiology (2). For papers published in American Society of Microbiology journals, the American Society of Microbiology does not require authors to obtain permission to include these papers in their thesis, provided the original work is properly cited. For all sheep in this study, differential cell counts were found to be within normal ranges except for sheep ED952 which had an elevated eosinophil count (19.2%) and sheep ED956 which had a high lymphocyte count (19.6%) which may indicate that these animals had a lung infection during this study. A larger version of Figure 1 can be found in **Appendix 7**. Data Set S1 can be found as Additional file 5_1.xlsx. Data Set S2 can be found as Additional file 5_2.xlsx. Data Set S3 can be found as Additional file 5_3.xlsx.

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

MICROBIAL ECOLOGY

Microbiota in Exhaled Breath Condensate and the Lung

Laura Glendinning, Steven Wright, Peter Tennant, Andrew C. Gill, David Collie, Gerry McLachlan

The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, Midlothian, United Kingdom

ABSTRACT The lung microbiota is commonly sampled using relatively invasive bronchoscopic procedures. Exhaled breath condensate (EBC) collection potentially offers a less invasive alternative for lung microbiota sampling. We compared lung microbiota samples retrieved by protected specimen brushings (PSB) and exhaled breath condensate collection. We also sought to assess whether aerosolized antibiotic treatment would influence the lung microbiota and whether this change could be detected in EBC. EBC was collected from 6 conscious sheep and then from the same anesthetized sheep during mechanical ventilation. Following the latter EBC collection, PSB samples were collected from separate sites within each sheep lung. On the subsequent day, each sheep was then treated with nebulized colistimethate sodium. Two days after nebulization, EBC and PSB samples were again collected. Bacterial DNA was quantified using 16S rRNA gene quantitative PCR. The V2-V3 region of the 16S rRNA gene was amplified by PCR and sequenced using Illumina MiSeq. Quality control and operational taxonomic unit (OTU) clustering were performed with mothur. The EBC samples contained significantly less bacterial DNA than the PSB samples. The EBC samples from anesthetized animals clustered separately by their bacterial community compositions in comparison to the PSB samples, and 37 bacterial OTUs were identified as differentially abundant between the two sample types. Despite only low concentrations of colistin being detected in bronchoalveolar lavage fluid, PSB samples were found to differ by their bacterial compositions before and after colistimethate sodium treatment. Our findings indicate that microbiota in EBC samples and PSB samples are not equivalent.

IMPORTANCE Sampling of the lung microbiota usually necessitates performing bronchoscopic procedures that involve a hospital visit for human participants and the use of trained staff. The inconvenience and perceived discomfort of participating in this kind of research may deter healthy volunteers and may not be a safe option for patients with advanced lung disease. This study set out to evaluate a less invasive method for collecting lung microbiota samples by comparing samples taken via protected specimen brushings (PSB) to those taken via exhaled breath condensate (EBC) collection. We found that there was less bacterial DNA in EBC samples compared with that in PSB samples and that there were differences between the bacterial communities in the two sample types. We conclude that while EBC and PSB samples do not produce equivalent microbiota samples, the study of the EBC microbiota may still be of interest.

KEYWORDS 16S rRNA, EBC, colistimethate sodium, colistin, lung microbiota, respiratory microbiota, sheep

Received 3 March 2017 **Accepted** 4 April 2017

Accepted manuscript posted online 7 April 2017

Citation Glendinning L, Wright S, Tennant P, Gill AC, Collie D, McLachlan G. 2017. Microbiota in exhaled breath condensate and the lung. Appl Environ Microbiol 83:e00515-17. https://doi.org/10.1128/AEM.00515-17.

Editor Harold L. Drake, University of Bayreuth

Address correspondence to Laura Glendinning, Laura.Glendinning@roslin.ed.ac.uk.

The study of the lung microbiota is a relatively new field in comparison to other areas of microbiota research. Although an increasing number of studies are linking changes in the composition of the lung bacterial communities to various disease states,

including allergies, autoimmune disorders, and inflammatory and infectious diseases (1), the protocols for studying the lung microbiota are not standardized, making comparisons between studies difficult.

One issue with studying the lung microbiota is the invasiveness of the sampling techniques; the most common techniques are bronchoalveolar lavage (BAL) and the collection of protected specimen brushings (PSB), both of which require the subject to undergo bronchoscopy. The inconvenience and fear of complications associated with bronchoscopic procedures can result in healthy and/or diseased individuals declining to take part in studies (2), leading to a reduction in the potential pool of volunteers for lung microbiota studies. It is also currently unknown whether these sampling methods themselves can lead to changes in the lung microbiota.

Exhaled breath condensate (EBC) collection could potentially be a less invasive method for taking lung microbiota samples. This method involves condensing exhaled vapor into a liquid and has previously been used to study exhaled bacteria, viruses, and fungi (3–8). However, there have been no studies using 16S rRNA gene sequencing to compare the bacteria found in EBC samples to those found in samples taken directly from the lungs. Therefore, it is not known whether it can be used as a surrogate for more-invasive sampling techniques. We sought to assess the feasibility of using EBC in sheep to study the lung microbiota composition. We have previously used sheep as a model for studying the lung microbiota (9, 10) due to the anatomical and immunological similarity of their lungs to those of humans (11–13). In this study, we compared EBC samples collected from conscious sheep and from the same sheep under anesthesia to PSB samples taken from four spatially disparate sites within the lungs.

We then extended this to address whether EBC analysis has the capacity to detect changes in bacterial community compositions by attempting to directly manipulate the lung microbiota with an inhaled antibiotic (colistimethate sodium [CMS], which is active against Gram-negative bacteria). In a previous study, we examined the effect of intravenous CMS on the lung microbiota (9). While we did identify changes in the lung microbiota composition, the longer-term systemic antibiotic treatment used in that study also likely affected the gut microbial populations. Immunological links between gut and lung immunities, the gut-lung axis, raise the possibility that such changes may have indirectly influenced the lung microbiota (14). In this study, we delivered nebulized CMS, since this has been shown to lead to lower colistin plasma concentrations than injected CMS (15), enabling us to discern the direct effect of antibiotic treatment on respiratory bacterial communities.

A far greater quantity of bacterial DNA was isolated from PSB samples relative to EBC samples. We found that while there was some overlap between the types of bacteria found in these samples, EBC samples did cluster separately from PSB samples by their bacterial community compositions. Lastly, despite our antibiotic treatment regime only producing low concentrations of colistin in the lung epithelial lining fluids (the prodrug CMS is hydrolyzed *in vivo* to the active form of the drug, colistin), significant differences in community compositions were found between PSB samples derived pre- and posttreatment.

RESULTS

Analysis of sequence quality and controls. DNA was extracted from respiratory samples and controls, and the V2-V3 variable regions of the 16S rRNA gene were amplified by PCR and then sequenced. After forming contigs from forward and reverse reads, various quality control steps were undertaken, which reduced the total sequence numbers by 25.8%. The lowest Good's coverage estimate value among the samples was 0.996, indicating that at least 99.6% of the bacteria in this sample were identified. The sequence error rate was 0.18% and the average number ($\pm$ standard deviation [SD]) of reads per sample was 39,195 $\pm$ 11,535. In total, 867 operational taxonomic units (OTUs) were identified.

The Human Microbiome Project mock community HM-783D, containing the 16S rRNA genes of 20 bacterial species in staggered quantities and fixed ratios (1,000 to

1,000,000 copies per organism per μl), was processed alongside the samples. Some biases were identified (see Data Set S2 in the supplemental material). Three species were incorrectly identified at the species level (*Acinetobacter baumannii* was misidentified as *Acinetobacter rhizosphaerae*, *Clostridium beijerinckii* was misidentified as *Clostridium butyricum*, and *Neisseria meningitidis* was misidentified as *Neisseria cinerea*). Two of the bacterial species which were present in low numbers in the original community were not identified at any taxonomic level, namely, *Actinomyces odontolyticus* and *Bacteroides vulgatus*. Their absence is likely due to the fact that they were in low abundance rather than the inability of our protocol to amplify and identify them, as they have previously been identified using the same protocol on a nonstaggered version of the same mock community (10). We were also previously able to identify *Enterococcus faecalis* at the genus level, whereas in this study, it could not be identified except potentially as OTU 10, *Bacilli* (class). This discrepancy, combined with the fact that *E. faecalis* is in low abundance in the staggered mock community, leads us to believe that this identification is incorrect.

As lung bacteria are in low abundance, lung samples are at a particular risk for contamination by bacterial DNA originating from DNA extraction kit reagents. Therefore, as well as mock community controls, DNA extraction kit reagent controls were produced. DNA was extracted from samples in four batches and a reagent control was included with every batch. The bacterial OTUs identified in the extraction kit controls did not occur consistently in samples from the same batch (Fig. 1). Samples were clustered by DNA extraction batch ($P < 0.001$ by analysis of molecular variance [AMOVA]), and 30 OTUs were found to be indicative of specific batches (see Data Set S3). However, when these OTUs were removed from the data set, samples still clustered by extraction batch ($P = 0.014$ by AMOVA), indicating that clustering was not entirely due to the presence of these OTUs. It is possible that some of these OTUs may be found naturally within the sheep respiratory system (e.g., *Micrococcus luteus*, a common colonizer of the human upper respiratory tract). Therefore, we decided not to remove these OTUs from our data set. Since samples were randomly assigned to extraction batches, clustering by batch would be unlikely to lead to false-positive statistical results. However, there is the possibility that the presence of contaminating organisms may increase the heterogeneity and thereby also increase stochastic noise.

In controls, the most abundant OTUs on average were as follows: *Corynebacterium*, 14.4%; *Enterobacteriaceae*, 10.9%; and *Intrasporangiaceae*, 3.6% in PSB controls and *Burkholderia*, 14.0%; *Neisseriaceae*, 10.5%; and *Aggregatibacter*, 7.7% in DNA extraction kit reagent controls. The most abundant OTUs (on average) in the different sample types were as follows: *Staphylococcus equorum*, 10.7%; *Mannheimia*, 6.5%; and *Staphylococcus sciuri*, 5.6% in PSB samples; *Staphylococcus equorum*, 5.5%; *Neisseriaceae*, 4.7%; and *Paracoccus*, 4.3% in EBC samples from conscious sheep (cons); and *Staphylococcus equorum*, 5.1%; *Staphylococcus epidermidis*, 3.7%; and *Peptostreptococcus anaerobius*, 3.2% in EBC samples from anesthetized sheep (anaes).

PSB samples contain more bacterial DNA than EBC samples. The V3 region of the 16S rRNA gene was quantified in our samples using quantitative PCR (qPCR). On average, PSB samples contained $1.53 \times 10^{-5} \pm 2.96 \times 10^{-5}$ ng/μl (mean ± SD) bacterial DNA (34,200 ± 66,100 16S copy numbers/μl), while EBC samples from conscious and anesthetized sheep contained $4.28 \times 10^{-7} \pm 5.34 \times 10^{-7}$ ng/μl (955 ± 1,190 16S copy numbers/μl) and $2.38 \times 10^{-7} \pm 7.12 \times 10^{-8}$ ng/μl (531 ± 159 16S copy numbers/μl), respectively (Fig. 2). DNA extraction kit reagent controls contained $1.82 \times 10^{-7} \pm 2.21 \times 10^{-8}$ ng/μl (406 ± 49 16S copy numbers/μl), while PSB controls and qPCR water controls contained $1.84 \times 10^{-7} \pm 1.05 \times 10^{-8}$ ng/μl (411 ± 23 16S copy numbers/μl) and $1.98 \times 10^{-7} \pm 2.06 \times 10^{-8}$ ng/μl (442 ± 46 16S copy numbers/μl), respectively.

All respiratory samples contained significantly more DNA than the controls ($P < 0.005$ for all sample types by Mann-Whitney U test). EBC samples from conscious and anesthetized animals did not contain significantly different quantities of DNA ($P =$

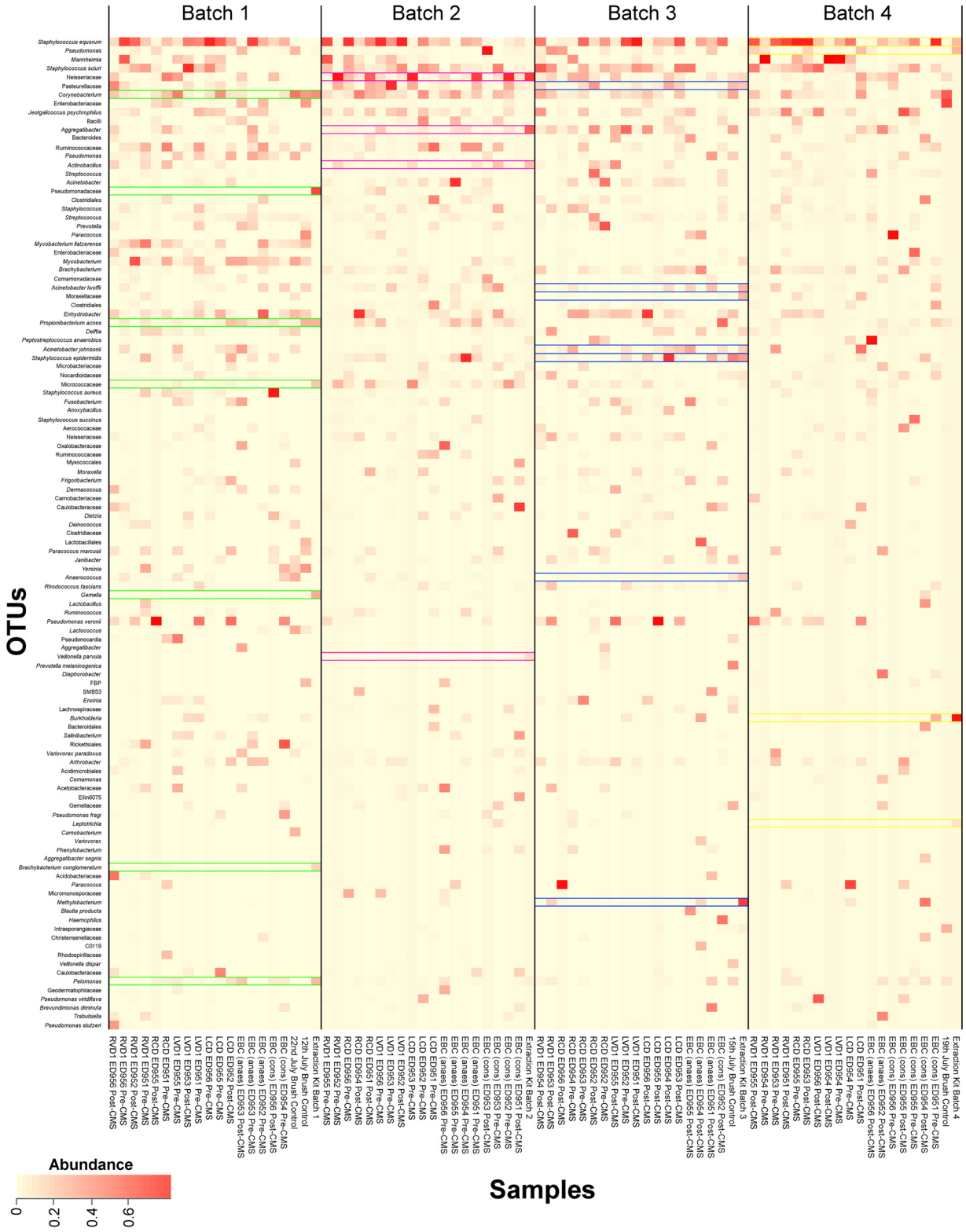


FIG 1 Heatmap showing samples grouped by batch based on the time DNA was extracted from the samples. Bacterial OTUs were included where they had an abundance of ≥5% in at least one sample. OTUs which were ≥5% abundant in a DNA extraction kit reagent control are indicated by color (batch 1, green; batch 2, pink; batch 3, blue; and batch 4, yellow). DNA extraction kit reagent controls are labeled as Extraction Kit Batch 1 to 4. EBC samples from conscious and from anesthetized sheep are labeled EBC (cons) and EBC (anaes), respectively. OTUs which were >5% abundant in an extraction kit control do not consistently appear in all samples in the same batch.

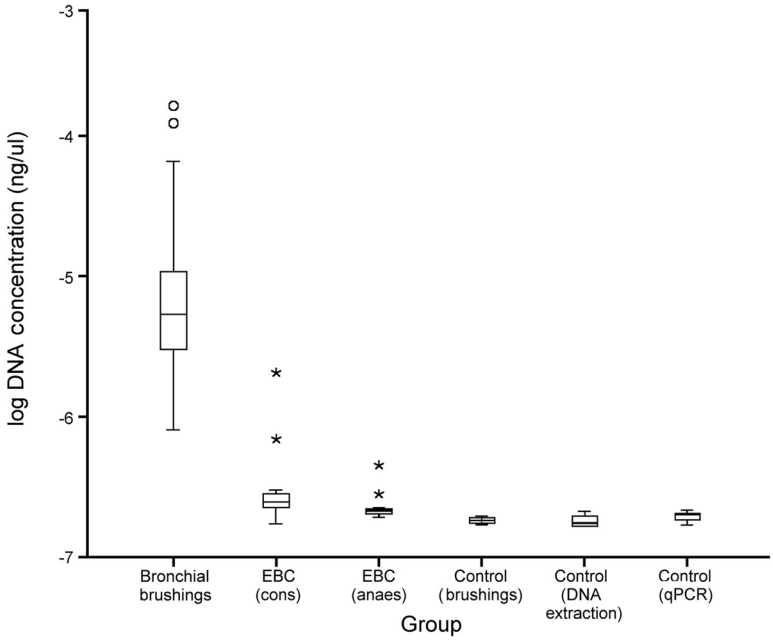


FIG 2 Boxplot showing the log 16S rRNA gene concentrations found in sheep respiratory samples (EBC samples from conscious and anesthetized animals and PSB samples) and controls (protected specimen brushes, DNA extraction kit reagents and qPCR reagents). Outliers were defined by SPSS as either "out" values (circles) or "extreme" values (stars). PSB samples contained significantly more bacterial DNA ($P < 0.005$) than any other respiratory sample type or control.

0.182 by Wilcoxon signed-rank test); however, PSB samples contained significantly more DNA than both EBC (cons) ($P = 0.002$ by Wilcoxon signed-rank test) and EBC (anaes) ($P = 0.002$ Wilcoxon signed-rank test) samples.

No significant clustering of EBC by sampling method. Since EBC samples from conscious sheep would be expected to include more bacteria from the upper respiratory tract than EBC samples from anesthetized sheep, it was expected that these two groups of samples would cluster separately from one another. However, no significantly separate clustering was observed ($P = 0.994$ by AMOVA). Despite this lack of separate clustering, EBC samples taken from the same sheep while it was conscious or anesthetized did not contain the same bacterial communities, as can be observed in Fig. 3.

The richness and the diversity of bacterial communities were not significantly different between the two groups ($P = 0.583$ and $P = 0.595$, respectively, by Wilcoxon signed-rank test). When examined using Metastats, there were significant differences in the quantities of several OTUs between these groups, but all of these OTUs were present at low abundances (<1% abundant on average in each group).

PSB samples and EBC (anaes) samples cluster separately by their bacterial communities. We next investigated whether PSB and EBC samples contained equivalent bacterial communities. We compared PSB samples with EBC (anaes) samples as we hypothesized that these would be less likely to be contaminated by upper respiratory tract microbes than EBC (cons) samples. As well as containing a larger quantity of bacterial DNA, PSB samples also contained bacterial communities that were significantly different from those of the EBC (anaes) samples ($P = 0.011$ by AMOVA) (Fig. 4). This may be explained by the difference in variation between the two groups ($P = 0.026$ by homogeneity of molecular variance [HOMOVA]). Bacterial communities from PSB samples were also found to be significantly richer ($P = 0.006$ by Wilcoxon signed-rank test), but there was no significant difference in diversity ($P = 0.48$ by Wilcoxon signed-rank test). One OTU designated *Pseudomonas veronii*, which was the 4th most abundant OTU in PSB samples, was found to be significantly more abundant in PSB samples (PSB samples [mean ± SD], 3.9% ± 1.3%; EBC, only one sequence read

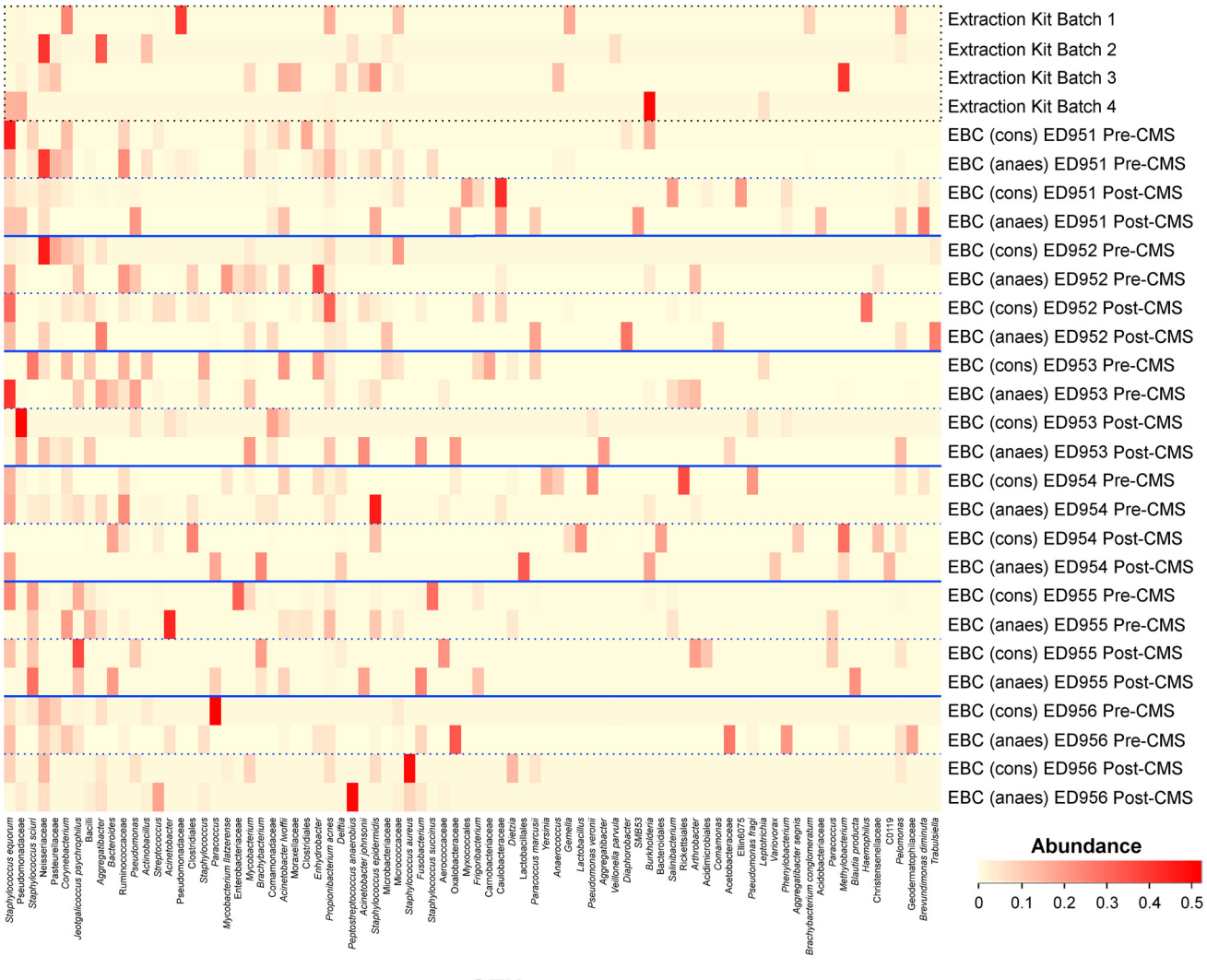


FIG 3 Heatmap showing EBC samples grouped by sheep and time point. DNA extraction kit reagent controls are labeled as Extraction Kit Batch 1 to 4. EBC samples from conscious and anesthetized sheep are labeled EBC (cons) and EBC (anaes), respectively. Bacterial OTUs were included where they had an abundance of ≥5% in at least one sample. As can be observed, EBC samples taken from the same sheep when it was conscious and when it was anesthetized did not necessarily contain the same bacterial OTUs.

found in one sample; Metastats q value = 0.046). The *P. veronii* OTU was not found in any of the PSB controls, indicating that its presence is not likely due to contamination. This indicates that the EBC samples do not simply contain a subset of the most abundant OTUs from PSB samples. An additional 36 low-abundance OTUs (<1% abundant on average in either group) were found to be significantly different between the two groups by Metastats.

We considered that since EBC (anaes) samples contained far less bacterial DNA than PSB samples, they may have been more affected by contamination and this may be why these sample types clustered separately. However, the five most abundant OTUs found in DNA extraction kit reagent controls (*Burkholderia*, *Neisseriaceae*, *Aggregatibacter*, *Pseudomonadaceae*, and *Methylobacterium*) were not found to be significantly differently represented between PSB samples and EBC (anaes) samples (Metastats q value = 1). Therefore, it seems unlikely that the separate clustering of these groups was due merely to the increased effect of contamination on EBC (anaes) samples.

Changes in the bacterial communities found in respiratory samples before and after CMS treatment. For both EBC (cons) and EBC (anaes) samples, pre- and post-treatment samples did not differ significantly by bacterial community structure (P =

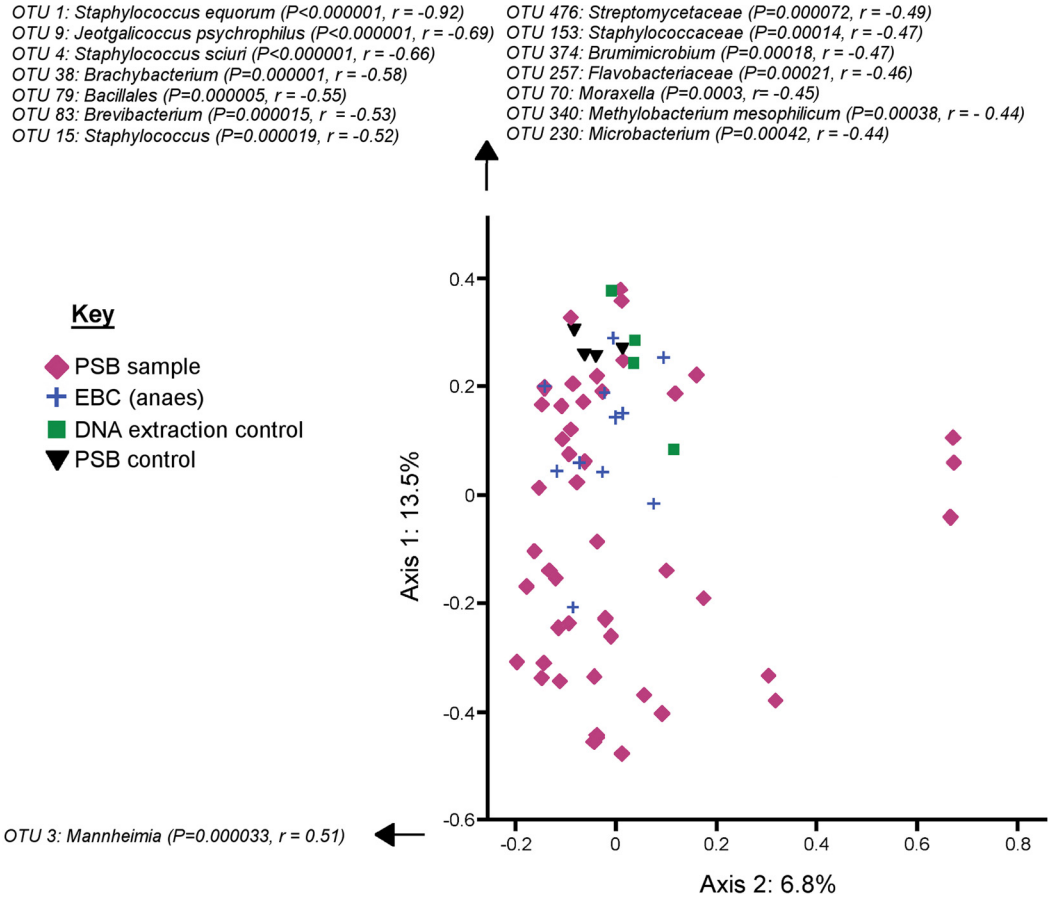


FIG 4 PCoA graph showing the significantly separate clustering of EBC (anaes) and PSB samples from sheep ($P = 0.011$ by AMOVA), which may be due to the difference in variation between the two sample types ($P = 0.026$ by HOMOVA). The OTUs which most contributed to samples moving in a positive or negative direction along either axis and which had P values of $<$ 0.00058 (defined by Bonferroni's correction as 0.5 divided by the total number of OTUs), according to the corr.axes command within mothur, are listed. As this graph is only representative of 20.3% of the total variability present between samples, caution should be taken when interpreting how clustered the sample groups appear.

0.449 and $P = 0.094$, respectively, by AMOVA). However, the bacterial communities found in PSB samples were found to be significantly different pre- and posttreatment ($P = 0.014$ by AMOVA) (Fig. 5). This significantly separate clustering was not merely due to differences in variation between the two groups ($P = 0.87$ by HOMOVA). The OTU *P. veronii* was increased in posttreatment samples (pretreatment [mean ± SD], 0.74% ± 0.39%; posttreatment, 7.1% ± 2.4%; Metastats q value = 0.043), and a further 97 low-abundance (<0.1%) OTUs were found to significantly differentiate pre- and posttreatment samples.

Using the Wilcoxon signed-rank test, it was found that the concentrations of DNA in respiratory samples before and after CMS treatment did not differ significantly: PSB samples, $P = 0.689$; EBC (cons) samples, $P = 0.345$; and EBC (anaes) samples, $P = 0.248$. The concentrations of colistin A identified in sheep lungs are shown in Table 1.

DISCUSSION

In this study, we sought to identify whether invasive lung microbiota sampling techniques could be replaced by a less invasive method. We compared the quantities of bacterial DNA and the bacterial communities from samples taken by PSB and EBC collection in six sheep at two sampling points. EBC was collected from both conscious and anesthetized animals. During mechanical ventilation, the animals were intubated, meaning that the exhaled breath collected was derived only from the lower respiratory tract. By comparing these samples to those taken from conscious animals, it should be

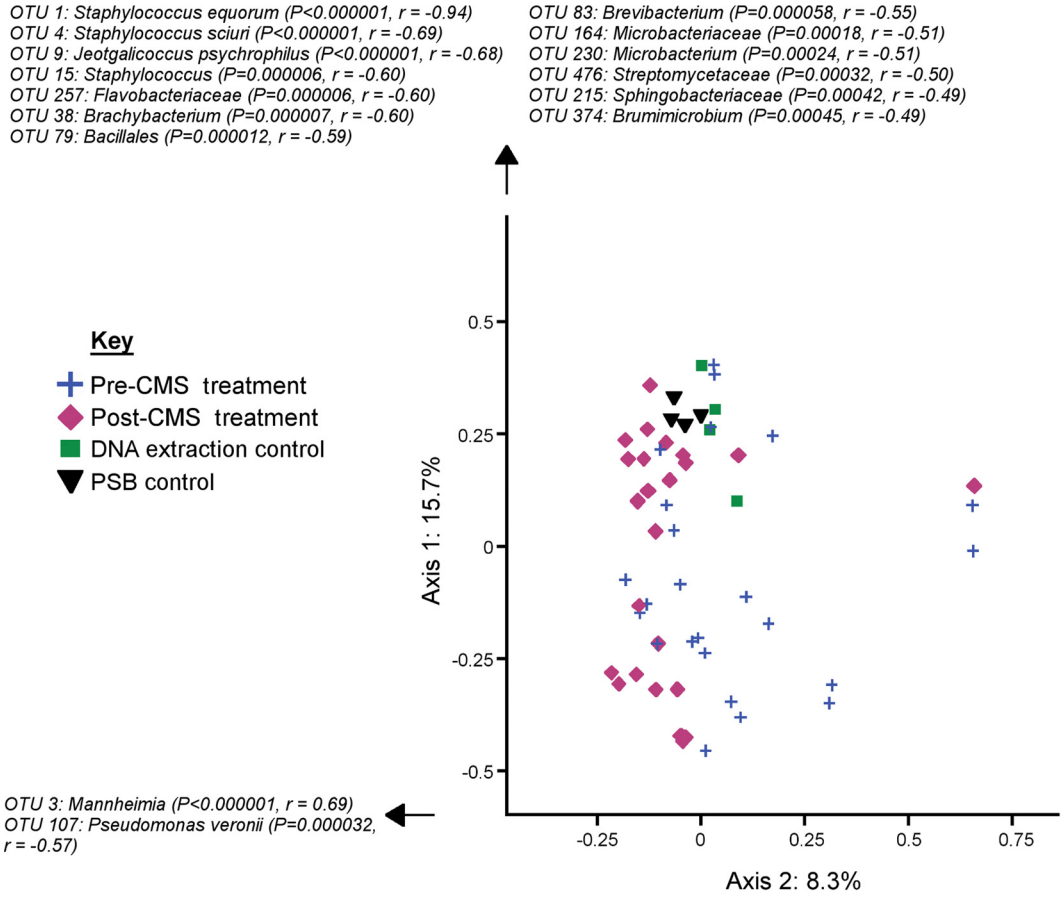


FIG 5 PCoA graph showing the significantly separate clustering of PSB samples from sheep before and after CMS treatment ($P = 0.014$ by AMOVA). The OTUs which most contributed to samples moving in a positive or negative direction along either axis and which had P values of < 0.00058 (defined by Bonferroni's correction as 0.5 divided by the total number of OTUs), according to the corr.axes command within mothur, are listed. As this graph is only representative of 24% of the total variability present between samples, caution should be taken when interpreting how clustered the sample groups appear.

possible to analyze the extent of contamination by bacteria from the upper respiratory tract in EBC (cons) samples. We found that EBC samples contained significantly less bacterial DNA than the PSB samples and that PSB samples clustered separately from EBC (anaes) samples by the composition of their bacterial communities. EBC (anaes) and EBC (cons) samples did not cluster separately from one another.

Studies examining the utility of EBC for identifying lung-colonizing microorganisms have shown variable results. A study comparing EBC and sputum samples from asthma patients showed a 100% overlap in the culturable fungi identified between the two sample types (5), and a study examining the bacterial pathogens cultured from BAL and EBC samples in patients with ventilator-associated pneumonia showed a high concordance between the two sampling methods (16). In comparison, when PCR assays for 10

TABLE 1 Colistin A concentrations in sheep epithelial lining fluid

Sheep	Colistin A concn[a] (ng/μl)	Dilution of epithelial lining fluid in BAL	Mean colistin A concn corrected for dilution (ng/μl)
ED951	0.346 ± 0.056	5.45	1.89
ED952	0.320 ± 0.034	4.18	1.34
ED953	0.290 ± 0.061	6.45	1.87
ED954	1.549 ± 0.251	15.75	24.40
ED955	0.625 ± 0.159	11.43	7.15
ED956	0.222 ± 0.017	29.5	6.56

[a]Values are the means ± SD. Colistin B values were too low to be calculated accurately.

common respiratory pathogens were performed on EBC and sputum samples from chronic obstructive pulmonary disease patients, the results were found not to correlate well (17). EBC collection has also previously been found to be inefficient for detecting *Mycobacterium tuberculosis* (18), influenza viruses (19), and the common cystic fibrosis pathogens *Pseudomonas aeruginosa* and *Burkholderia cepacia* (20).

Some concerns have been raised about the use of EBC in respiratory research, since the epithelial lining fluid contained in these samples is often variable and is very highly diluted with water vapor (21). This dilution could explain the far lower concentrations of bacterial DNA we identified in EBC samples in comparison to those from PSB samples. It is also likely that PSB would be more efficient for sampling biofilms adhered to the lung mucosa, which could explain some of the differences observed between the two sample types. The difference between the bacterial communities found in PSB and EBC samples may also be partially explained by how EBC is formed. The exact origin of EBC is still under debate, but it has been suggested that differences observed between BAL and EBC samples could be explained by the fact that different compartments of the lung are sampled (22). While it might be assumed that EBC would be derived from both the central and peripheral airway compartments, which would perhaps explain the differences between these samples and PSB samples, Bondesson et al. concluded that the majority of EBC is in fact derived from the central airways (23). Without a better understanding of how EBC is formed and what influences its composition, we are unable to account for the differences we observed between the two sampling types.

Despite the fact that the concentrations of colistin found in the lungs were quite low after nebulized CMS treatment, a significant difference was observed in the bacterial communities from PSB samples pre- and posttreatment. In a previous study, we found that the relative proportion of Gram-negative bacteria in the lung microbiota (excluding *Pseudomonadales*) was reduced after injected CMS treatment (9). However, members of *Pseudomonadales* generally increased in relative abundance or remained stable after treatment. Therefore, it is interesting to note that while in this study we did not find a significant reduction in the abundance of Gram-negative bacteria in PSB samples (data not shown), an OTU belonging to *Pseudomonadales* (*P. veronii*) was significantly increased in these samples after CMS treatment. It is possible that, even at low concentrations, the colistin may have had some effect on the lung bacteria or that the sampling strategy may itself in some way lead to changes in the lung microbiota, but at the moment, this is purely speculative. All samples were randomized prior to DNA extraction and PCR amplification; therefore, the observed differences were not due to samples from one time point being processed separately from those from the other time point.

In conclusion, the differences we observed between PSB samples and EBC samples lead us to not recommend using EBC collection as a replacement for more-invasive lung sampling techniques. However, the EBC microbiota may still be an interesting avenue of study despite the fact that the small quantities of bacterial DNA in these samples leave them more vulnerable to contamination, and any future studies would have to be designed with this in mind.

MATERIALS AND METHODS

Animals. Six commercially sourced, castrated male Suffolk-cross sheep aged 14 months were used in this study. All animal experiments were approved by the Roslin Institute Animal Welfare and Ethics Committee and were subject to the Animals (Scientific Procedures) Act of 1986. Sheep had previously been housed outdoors as part of a large flock but were moved indoors before the study and remained indoors until the study end. Sheep were separated into two pens sharing the same airspace. One pen contained sheep ED951, ED952, and ED953, while the other contained sheep ED954, ED955, and ED956. The rectal temperatures and weights of all animals were taken prior to the initial respiratory tract sampling. The animals weighed on average ($\pm$ SD) 49.2 $\pm$ 3.4 kg and the rectal temperatures were measured as 38.9 $\pm$ 0.89°C.

Experimental design. Conscious animals were confined in a yoke head-restraint holding crate, and EBC was collected for 10 min using an RTubeVENT with a cooling sleeve (Respiratory Research, Charlottesville, VA, USA) attached to a face mask. The sheep inhaled through a one-way inspiratory valve and expired through the RTubeVENT (Fig. 6). The exhaled breath condensate samples from conscious

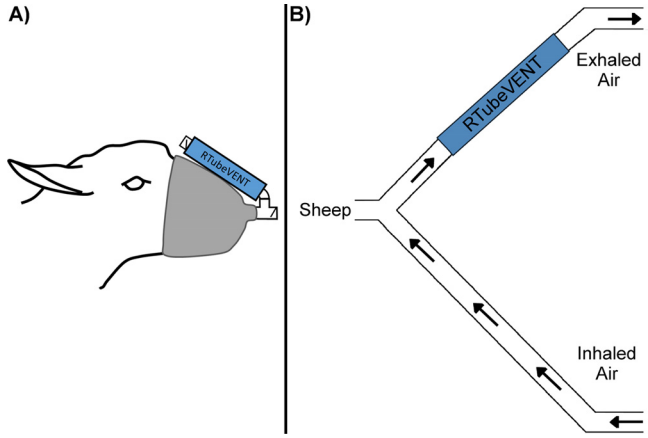


FIG 6 (A) EBC was collected from conscious animals while they were restrained in a yoke head-restraint holding crate. A face mask was attached and sheep inhaled through a short tube with an inlet valve and exhaled through an RTubeVENT. (B) EBC was collected from anesthetized mechanically ventilated animals by placing the RTubeVENT in-line with the expiratory limb of the ventilator, near the sheep's head.

sheep (EBC [cons]) were transferred from the RTubeVENT into Eppendorf tubes according to the manufacturer's instructions and were frozen on dry ice within an hour of collection.

The sheep were then anesthetized (3 to 5 h later) according to a procedure that has previously been described (24). Bronchoscopy was performed using an endotracheal tube. During anesthesia, EBC samples were collected for 10 min by incorporating an RTubeVENT within the expiratory limb of the anesthetic circuit (Fig. 6). The condensate was again transferred into Eppendorf tubes. The exhaled breath condensate samples from the anesthetized sheep (EBC [anaes]) were frozen on dry ice within an hour of collection. PSB samples (disposable microbiology brush; Conmed, Utica, NY, USA) were taken from the left ventral diaphragmatic 1 (LVD1), right ventral diaphragmatic 1 (RVD1), right caudal diaphragmatic (RCD), and left caudal diaphragmatic (LCD) lung segments (Fig. 7). Brushes were cut into phosphate-buffered saline (PBS; Sigma-Aldrich, Irvine, UK) for storage. For each sampling day, an unused protected specimen brush was cut into PBS to act as a control.

Eighteen hours after the recovery from anesthesia, sheep were administered 2,000,000 IU of CMS in 4 ml distilled water by inhalation (Colomycin for injection; Forest Laboratories UK Ltd., Dartford, UK). Restraint of the conscious sheep was as described above, and the CMS was delivered using a face mask connected via the inspiratory limb to an eFlow rapid nebulizer (PARI Respiratory Equipment Inc., Midlothian, VA, USA). This treatment was repeated 6 h later. Two days after the first CMS dose was administered, EBC (cons), EBC (anaes), and PSB samples were again collected as described above. Sheep were killed by barbiturate overdose and exsanguination and blood samples were collected. Blood was centrifuged at 2,500 $\times$ *g* for 5 min and the serum was removed and frozen on dry ice. Immediately postmortem, 20-ml aliquots of PBS were used to collect BAL fluid. The urea concentrations in plasma and BAL fluid were used to calculate the dilution factor of lung epithelial lining fluid in BAL fluid (25).

Quantitation of colistin in BAL fluid of sheep. BAL fluid was centrifuged at 1,400 $\times$ *g* for 5 min to remove cells prior to colistin quantification. The quantitation of colistin in ovine BAL fluid essentially follows the method previously published by Marchand et al. (26). Briefly, colistin sulfate (item no. 17584 [mixture of A and B isoforms]; Cayman Chemicals, Ann Arbor, MI, USA) was dissolved in H_2O to 1 mg/ml and a series of 7 calibrant solutions were created by diluting the stock solution into blank BAL fluid to cover the range from 100 to 0.07 μg/ml. Polymyxin B (Sigma-Aldrich, Irvine, UK) was used as an internal standard and was dissolved in water to 300 μg/ml. Two microliters of internal standard was added to 200 μl of each of the calibrant solutions and to 200 μl of each of the test samples. Eight hundred microliters of a solution of H_2O and 0.1% (vol/vol) formic acid was added to each of the samples/calibrants, and each was partially purified by binding to a DSC-18 SPE cartridge (Sigma-Aldrich, Irvine, UK) and eluted with 400 μl methanol (MeOH) and 0.1% (vol/vol) formic acid. The eluted fractions were dried under vacuum and reconstituted in 50 μl of H_2O and 0.1% (vol/vol) formic acid for subsequent analysis.

All calibrants and samples were centrifuged at 13,000 $\times$ *g* for 5 min to pellet any precipitate and then were analyzed by online liquid chromatography tandem-mass spectrometry (LC-MS/MS) in duplicate. Aliquots of 5 μl were injected into an Ace Ultracore 2.5 SuperC18 high-performance liquid chromatography (HPLC) column (75 mm by 2.1 mm) preequilibrated with 98% (vol/vol) buffer A, where HPLC buffer A was H_2O with 0.1% (vol/vol) formic acid and 0.01% (vol/vol) trifluoroacetic acid, while HPLC buffer B was acetonitrile with 0.1% (vol/vol) formic acid and 0.01% (vol/vol) trifluoroacetic acid. The HPLC separation was developed by the following steps: from 2% buffer B at 0 min to 18% buffer B at 1 min, 22% buffer B at 3.5 min, 100% buffer B at 4 min, 100% buffer B at 5 min, and returning to 2% buffer B at 6 min for 5 min to reequilibrate. The flow rate was 200 μl/min and the eluent was passed directly to

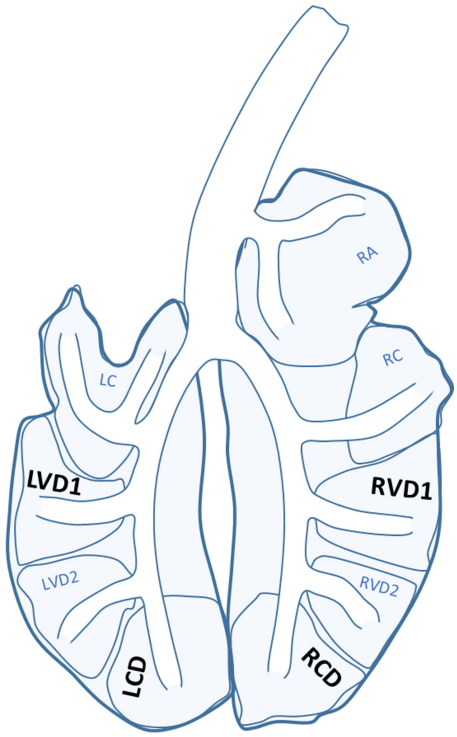


FIG 7 Diagram of the sheep lung. PSB samples were taken before and after colistimethate sodium treatment from the right ventral diaphragmatic 1 (RVD1), left ventral diaphragmatic 1 (LVD1), right caudal diaphragmatic (RCD), and left caudal diaphragmatic (LCD) lung segments. RC, right cardiac; RA, right apical; LC, left cardiac; LVD2, left ventral diaphragmatic 2; RVD2, right ventral diaphragmatic 2. Adapted from reference 24.

the electrospray source of an Amazon ETD ion trap mass spectrometer (Bruker, Billerica, MA, USA) operated in positive-ion mode. The mass spectrometer was operated under multiple reaction monitoring conditions, using parent ions of 578.3, 585.3, and 602.3 (representing the double-charged ion of colistin B, colistin A, and polymyxin B, respectively), fragmentation amplitudes of 0.8, and cutoffs of 140 in each case. Calibration curves and colistin concentrations were calculated by Bruker's proprietary software QuantAnalysis using the following reporter ions: 526.3, 535.3, 567.3, and 576.3 (colistin A); 519.3, 528.3, 560.800, and 569.3 (colistin B); and 543.300, 552.300, 584.300, and 593.3 (polymyxin B).

DNA extraction. DNA extraction was carried out as described previously (10) using the Mo Bio PowerSoil DNA isolation kit (Mo Bio Laboratories Inc., Carlsbad, CA, USA). All DNA extractions were carried out using extraction kits from the same lot, as the contamination present in different lots of the same make of kit has been shown not to be consistent (27). Samples were randomly assigned to one of four DNA extraction batches, and for each of these batches, an extraction kit reagent-only control was produced (sample groupings can be found in Data Set S1 in the supplemental material).

16S rRNA gene amplification and sequencing. The V2-V3 variable regions of the 16S gene were amplified as described previously (10). A nested PCR protocol was used to decrease the potential bias introduced by the use of barcoded primers by only including primers with Illumina adaptor sequences and barcodes in the second PCR round (28). The first round used the V1-V4 primers 28F (5′-GAGTTTG ATCNTGGCTCAG-3′) and 805R (5′-GACTACCAGGGTATCTAATC-3′) and the second round used the V2-V3 primers 104F (5′-GGCGVACGGGTGAGTAA-3′) and 519R (5′-GTNTTACNGCGGCKGCTG-3′) with Illumina adaptor sequences and barcodes (Data Set S1). The PCR conditions for the first round were 94°C for 2 min followed by 20 cycles of 94°C for 1 min, 55°C for 45 s and 72°C for 1.5 min, followed by 72°C for 20 min. The conditions for the second round were 98°C for 30 s followed by 20 cycles of 98°C for 10 s, 67°C for 30 s, and 72°C for 10 s, followed by 72°C for 2 min. Q5 High-fidelity 2× master mix (New England BioLabs, Ipswich, MA, USA) was used for all reactions. After each PCR round, amplicons were purified using the AMPure XP PCR purification system (Beckman Coulter, Brea, CA, USA). The Human Microbiome Project mock community HM-783D (obtained through BEI Resources, NIAID, NIH) also underwent PCR alongside samples and controls. The Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Hemel Hempstead, UK) was used to calculate the quantity of DNA in each sample, and then samples were pooled into a sequencing library. Sequencing was performed using Illumina MiSeq (Illumina, San Diego, CA, USA) producing 250-bp paired-end reads.

Bioinformatic and statistical analysis. Primers were removed with cutadapt (29) and sequences with greater than one base error per 10 bases were discarded. Quality control, taxonomic assignment, and OTU clustering were performed in mothur (30) as described previously (10). The data were

subsampled to the minimum number of sequence reads found in one of our samples (11,675). Except where stated, the following analyses were all performed within mothur.

Good's coverage values were calculated to estimate sample coverage (31). Distance matrices were constructed using Yue-Clayton theta values (32), and AMOVA was used to compare groups of samples by their bacterial composition (33). HOMOVA was used to compare groups by their variation (34). Principal-coordinate analysis (PCoA) graphs were constructed to visualize sample clustering. The mothur command corr.axes was used to correlate bacterial OTUs to the axes of the PCoA graphs using the Spearman's rank correlation coefficient (*r*). Bonferroni's correction was used to correct for multiple statistical tests. The inverse Simpson's index was employed to measure microbial diversity and the Chao 1 index was employed to measure richness. Metastats was used to identify OTUs which were significantly different between groups (35) except where more than two groups were compared, in which case indicator analysis was used (36).

To compare groups statistically when data were nonparametric, the Mann-Whitney U test was used if the groups were independent and the Wilcoxon signed-rank test was used when samples were related (performed in SPSS Statistics 21; IBM Analytics). Boxplots for qPCR data were constructed in SPSS. Heatmaps were constructed in R (version 3.2.2; R Foundation for Statistical Computing) using the packages gplots (37), heatplus (38), RColorBrewer (39), and Vegan (40).

qPCR. Quantification of the V3 region of the 16S rRNA gene was carried out using a previously described method (10). A standard curve was generated using DNA extracted from *Pseudomonas aeruginosa* strain PA0579 using 9 serial dilutions ranging from 14.2 ng/μl to 1.42 $\times$ 10^{-7} ng/μl (quantified by the Qubit dsDNA HS assay). The 0.142 ng/μl dilution served as a positive control for all qPCRs. The average threshold cycle (C_T) value of no-template controls was 28.7.

qPCR was performed using 1 μl of extracted DNA solution, the primers UniF340 (5′-ACTCCTACGG GAGGCAGCAGT-3′) and UniR514 (5′-ATTACCGCGGCTGCTGGC-3′) at a final concentration of 0.4 μM, and the LightCycler 480 SYBR green I master mix (Roche Applied Science, Mannheim, Germany). The qPCR run consisted of a preincubation step (50°C for 2 min and 95°C for 10 s), an amplification step (45 cycles of 95°C for 30 s and then 63°C for 30 s), and a melting cycle.

Accession number(s). Sequencing reads can be accessed under BioProject accession number PRJNA337937.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.00515-17.

SUPPLEMENTAL FILE 1, XLSX file, 0.1 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.1 MB.
SUPPLEMENTAL FILE 3, XLSX file, 0.1 MB.

ACKNOWLEDGMENTS

The following reagent was obtained through BEI Resources, NIAID, NIH, as part of the Human Microbiome Project: genomic DNA from microbial mock community B (staggered, low concentration), v5.2L, for 16S rRNA gene sequencing, HM-783D. We thank the staff of Dryden Animal Services for the care of our animals.

Funding for this study was supplied by the BBSRC (Biotechnology and Biological Sciences Research Council, UK) under the EASTBIO doctoral training program (grant number BB/J01446X/1) and, in part, by an Institute Strategic Programme grant (no. BB/J004332/1) also from the BBSRC.

The funders had no input on study design, data collection and interpretation, or the decision to submit this study for publication.

REFERENCES

1. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. 2016. The microbiome and the respiratory tract. Annu Rev Physiol 78:481–504. https://doi.org/10.1146/annurev-physiol-021115-105238.
2. Martinsen EMH, Leiten EO, Bakke PS, Eagan TML, Gronseth R. 2016. Participation in research bronchoscopy: a literature review. Eur Clin Respir J 3:29511. https://doi.org/10.3402/ecrj.v3.29511.
3. Xu Z, Shen F, Li X, Wu Y, Chen Q, Jie X, Yao M. 2012. Molecular and microscopic analysis of bacteria and viruses in exhaled breath collected using a simple impaction and condensing method. PLoS One 7:e41137. https://doi.org/10.1371/journal.pone.0041137.
4. Bhimji A, Singer LG, Kumar D, Humar A, Pavan R, Zhang H, Rotstein C, Keshavjee S, Mazzulli T, Husain S. 2016. Feasibility of detecting fungal DNA in exhaled breath condensate by the Luminex Multiplex xTAG fungal PCR assay in lung transplant recipients: a pilot study. J Heart Lung Transplant 35:S37. https://doi.org/10.1016/j.healun.2016.01.099.
5. Carpagnano GE, Malerba M, Lacedonia D, Susca A, Logrieco A, Carone M, Cotugno G, Palmiotti GA, Foschino-Barbaro MP. 2016. Analysis of the fungal microbiome in exhaled breath condensate of patients with asthma. Allergy Asthma Proc 37:41–46. https://doi.org/10.2500/aap.2016.37.3943.
6. Carpagnano GE, Lacedonia D, Palladino GP, Logrieco G, Crisetti E, Susca A, Logrieco A, Foschino-Barbaro MP. 2014. *Aspergillus* spp. colonization in exhaled breath condensate of lung cancer patients from Puglia region of Italy. BMC Pulm Med 14:22. https://doi.org/10.1186/1471-2466-14-22.
7. Carpagnano GE, Lacedonia D, Cotugno G, Depalo A, Palmiotti GA, Giuffreda E, Foschino Barbaro MP. 2015. Exhaled fungal microbiome in asthma: possible association with severity and control. Eur Respir J 46:PA4004. https://doi.org/10.1183/13993003.congress-2015.PA4004.
8. McDevitt JJ, Koutrakis P, Ferguson ST, Wolfson JM, Fabian MP, Martins M, Pantelic J, Milton DK. 2013. Development and performance evaluation of

an exhaled-breath bioaerosol collector for influenza virus. Aerosol Sci Technol 47:444–451. https://doi.org/10.1080/02786826.2012.762973.
9. Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, Doherty C, McLachlan G. 2015. Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate sodium. PLoS One 10:e0142097. https://doi.org/10.1371/journal.pone.0142097.
10. Glendinning L, Wright S, Pollock J, Tennant P, Collie D, McLachlan G. 2016. Variability of the sheep lung microbiota. Appl Environ Microbiol 82:3225–3238. https://doi.org/10.1128/AEM.00540-16.
11. Bouljihad M, Leipold HW. 1994. An ultrastructural study of pulmonary bronchiolar and alveolar epithelium in sheep. Zentralbl Veterinarmed A 41:573–586. https://doi.org/10.1111/j.1439-0442.1994.tb00125.x.
12. Griebel PJ, Entrican G, Rocchi M, Beskorwayne T, Davis WC. 2007. Cross-reactivity of mAbs to human CD antigens with sheep leukocytes. Vet Immunol Immunopathol 119:115–122. https://doi.org/10.1016/j.vetimm.2007.06.015.
13. Scheerlinck JPY, Snibson KJ, Bowles VM, Sutton P. 2008. Biomedical applications of sheep models: from asthma to vaccines. Trends Biotechnol 26:259–266. https://doi.org/10.1016/j.tibtech.2008.02.002.
14. Dickson RP, Morris A. 2017. Macrolides, inflammation and the lung microbiome: untangling the web of causality. Thorax 72:10–12. https://doi.org/10.1136/thoraxjnl-2016-209180.
15. Yapa SWS, Li J, Patel K, Wilson JW, Dooley MJ, George J, Clark D, Poole S, Williams E, Porter CJH, Nation RL, McIntosh MP. 2014. Pulmonary and systemic pharmacokinetics of inhaled and intravenous colistin methanesulfonate in cystic fibrosis patients: targeting advantage of inhalational administration. Antimicrob Agents Chemother 58:2570–2579. https://doi.org/10.1128/AAC.01705-13.
16. May AK, Brady JS, Romano-Keeler J, Drake WP, Norris PR, Jenkins JM, Isaacs RJ, Boczko EM. 2015. A pilot study of the noninvasive assessment of the lung microbiota as a potential tool for the early diagnosis of ventilator-associated pneumonia. Chest 147:1494–1502. https://doi.org/10.1378/chest.14-1687.
17. Zakharkina T, Koczulla AR, Mardanova O, Hattesohl A, Bals R. 2011. Detection of microorganisms in exhaled breath condensate during acute exacerbations of COPD. Respirology 16:932–938. https://doi.org/10.1111/j.1440-1843.2011.01977.x.
18. Jain R, Schriever CA, Danziger LH, Cho SH, Rubinstein I. 2007. The IS6110 repetitive DNA element of *Mycobacterium tuberculosis* is not detected in exhaled breath condensate of patients with active pulmonary tuberculosis. Respiration 74:329–333. https://doi.org/10.1159/000101786.
19. St George K, Fuschino ME, Mokhiber K, Triner W, Spivack SD. 2010. Exhaled breath condensate appears to be an unsuitable specimen type for the detection of influenza viruses with nucleic acid-based methods. J Virol Methods 163:144–146. https://doi.org/10.1016/j.jviromet.2009.08.019.
20. Vogelberg C, Hirsch T, Rosen-Wolff A, Kerkmann ML, Leupold W. 2003. *Pseudomonas aeruginosa* and *Burkholderia cepacia* cannot be detected by PCR in the breath condensate of patients with cystic fibrosis. Pediatr Pulmonol 36:348–352. https://doi.org/10.1002/ppul.10352.
21. Effros RM, Dunning MB, Biller J, Shaker R. 2004. The promise and perils of exhaled breath condensates. Am J Physiol Lung Cell Mol Physiol 287:L1073–L1080. https://doi.org/10.1152/ajplung.00069.2004.
22. Jackson AS, Sandrini A, Campbell C, Chow S, Thomas PS, Yates DH. 2007. Comparison of biomarkers in exhaled breath condensate and bronchoalveolar lavage. Am J Respir Crit Care Med 175:222–227. https://doi.org/10.1164/rccm.200601-107OC.
23. Bondesson E, Jansson LT, Bengtsson T, Wollmer P. 2009. Exhaled breath condensate-site and mechanisms of formation. J Breath Res 3:016005. https://doi.org/10.1088/1752-7155/3/1/016005.
24. Collie D, Govan J, Wright S, Thornton E, Tennant P, Smith S, Doherty C, McLachlan G. 2013. A lung segmental model of chronic *Pseudomonas* infection in sheep. PLoS One 8:e67677. https://doi.org/10.1371/journal.pone.0067677.
25. Rennard SI, Basset G, Lecossier D, Odonnell KM, Pinkston P, Martin PG, Crystal RG. 1986. Estimation of volume of epithelial lining fluid recovered by lavage using urea as marker of dilution. J Appl Physiol 60:532–538.
26. Marchand S, Gobin P, Brillault J, Baptista S, Adier C, Olivier J-C, Mimoz O, Couet W. 2010. Aerosol therapy with colistin methanesulfonate: a biopharmaceutical issue illustrated in rats. Antimicrob Agents Chemother 54:3702–3707. https://doi.org/10.1128/AAC.00411-10.
27. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12:87. https://doi.org/10.1186/s12915-014-0087-z.
28. Berry D, Ben Mahfoudh K, Wagner M, Loy A. 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl Environ Microbiol 77:7846–7849. https://doi.org/10.1128/AEM.05220-11.
29. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12. https://doi.org/10.14806/ej.17.1.200.
30. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541. https://doi.org/10.1128/AEM.01541-09.
31. Esty WW. 1986. The efficiency of Good's nonparametric coverage estimator. Ann Stat 14:1257–1260. https://doi.org/10.1214/aos/1176350066.
32. Yue JC, Clayton MK. 2005. A similarity measure based on species proportions. Commun Stat Theory Methods 34:2123–2131. https://doi.org/10.1080/STA-200066418.
33. Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.
34. Stewart CN, Excoffier L. 1996. Assessing population genetic structure and variability with RAPD data: application to *Vaccinium macrocarpon* (American Cranberry). J Evol Biol 9:153–171. https://doi.org/10.1046/j.1420-9101.1996.9020153.x.
35. Paulson J, Pop M, Bravo H. 2011. Metastats: an improved statistical method for analysis of metagenomic data. Genome Biol 12(Suppl 1):P17. https://doi.org/10.1186/gb-2011-12-S1-P17.
36. Dufrene M, Legendre P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecol Monogr 67:345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2.
37. Warnes G, Bolker B, Bonebakker L, Gentleman R, Liaw W, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. 2015. gplots: various R programming tools for plotting data, v2.17.0. http://CRAN.R-project.org/package=gplots.
38. Ploner A. 2015. Heatplus: Heatmaps with row and/or column covariates and colored clusters, v2.16.0. https://github.com/alexploner/Heatplus.
39. Neuwirth E. 2014. RColorBrewer: ColorBrewer palettes, v1.1-2. http://CRAN.R-project.org/package=RColorBrewer.
40. Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H. 2016. Vegan: community ecology package, v2.3-4/r2988. http://R-Forge.R-project.org/projects/vegan/.

5.3 Discussion

It is interesting to note that the concentrations of colistin achieved in this study were far lower than was achieved in the sheep described in Section 5.1. More doses were delivered to this sheep over a longer treatment period than were administered to the study sheep, which may explain this discrepancy. In future studies I would recommend using this more prolonged treatment regime in order to ensure that bactericidal levels of colistin are produced in the epithelial lining fluid.

We were able to demonstrate that samples taken directly from the lung were not equivalent to samples taken by EBC. In Chapters 3, 4 and 5 samples have been collected directly from the lung in separate sheep populations. It is possible that by comparing the results from these Chapters more general conclusions about the composition of the sheep lung microbiota across populations could be reached.

Chapter 6: Identifying Core Members of the Sheep Lung Microbiota

6.1 Introduction

During the previous results chapters I have explored the variability of the lung microbiota spatially and between sheep, compared upper and lower respiratory tract samples and examined the potential use of exhaled breath condensate as a non-invasive lung microbiota sampling method. During each of these chapters lung sampling methods were chosen which did not expose samples to potential contamination from the upper respiratory tract; in Chapters 3 and 5 this consisted of protected specimen brushings in live animals and in Chapter 4 it consisted of PBS lung washes of dead animals.

While the types of bacteria identified in these samples can begin to give us some idea of the composition of the sheep lung microbiota, none of these chapters individually are able to address whether there is a 'core' lung microbiota shared across sheep populations as each chapter contains samples from only one population of sheep. The differences observed between samples taken from the oropharynx and lower respiratory tracts of lambs in Chapter 4 would indicate that if a core sheep lung microbiota exists it is unlikely to be reminiscent of the oropharynx. If core bacterial OTUs can be identified in the sheep lung, it would be interesting to culture and better characterise these microbes, potentially leading to a better understanding of their significance and role in the lung.

By comparing the bacterial OTUs found in lung samples across my results chapters it may be possible to identify OTUs which are found in lung samples from all three sheep populations. One potential issue with this approach is that differences in the methods used between chapters might lead to differences in results, including the use of different lung sampling techniques. Therefore, in this chapter I first compare sequencing errors, taxonomic identifications, coverage and DNA concentrations in control and lung samples between results chapters. I also compare staggered and even mock communities, in order to identify any biases which may be present due to my PCR amplification and sequencing strategies. Finally, I compare lung samples to reagent controls across results chapters to create a list of potential core members of the sheep lung microbiota. Where lung samples are referred to in this chapter this refers only to samples taken directly from the lungs (lung brushings and lung washes) and does not include EBC or upper respiratory tract samples (see **Table 6.1** for the full list of samples included).

Table 6.1: Samples included in this chapter

Chapter	Sample types	Sample details
3	Protected specimen brushings	N=6 sheep (20 months old, 5 females and one castrated male): All brushings from all time-points. Total of 54 samples. N=1 sheep (36 months old, female): Brushings A3-A9 and A13-A19. Total of 14 samples.
4	Whole lung lavage	N=40 sheep (48.8 days (mean) ± 0.8 (SD) old, 20 females and 20 males. All lung fluid samples, totalling 40 samples.
5	Protected specimen brushings	N=6 sheep (14 months old, 6 castrated males). All protected specimen brushing samples, totalling 48 samples.

6.2 Results and discussion

6.2.1 Differences in error rate, taxonomic identification, coverage and sequence numbers across results chapters

Caution should be taken when comparing results across my thesis chapters as different methodologies were used in each chapter. While ideally I would have used the same sequencing platform and the same number of sequencing runs for every results chapter, this was unfortunately not possible. Differences caused by changes in sequencing methodology may make comparisons between chapters more difficult. When comparing sequence error rates, sequences per sample and the level of taxonomic depth to which sequences were classified, variation was observed between chapters (**Table 6.2**). While the sequencing error rates are similar for Chapters 3 and 4 (0.39% and 0.35%), the error rate for Chapter 5 is substantially reduced (0.18%). Chapter 5 also had the highest percentage reduction in sequences during quality control. The average number of sequences per sample was also far less in Chapter 5 compared to the other results Chapters, which would be expected to cause decreased coverage. The lowest average Good's coverage value was found in Chapter 5 (0.992 ± 0.00369 (mean ± SD), **Table 6.2**).

Some of the differences between chapters can be accounted for by a problem which occurred in 2015 with the Illumina MiSeq reagents which led to greatly decreased quality of longer reads (personal communication – Edinburgh Genomics). As a result, samples from between June 2015 (when the problem arose) and May 2016 (when improved sequencing kits became available) were sequenced using the Illumina HiSeq. Some of the samples included in Chapters 3 and 4 were sequenced during this period and some were sequenced prior to it. All sequencing for Chapter 5 was done post May

2016. The Illumina HiSeq produces ten times as many reads per run than the MiSeq which explains the differences in average read numbers per sample. It is possible that the improved sequencing kits are responsible for the decrease in the error rate observed in Chapter 5 but this is currently speculative.

Between chapters, a similar percentage of sequences were able to be assigned to the taxonomic level of family (Chapter 3: 95.6%, Chapter 4: 95.4%, Chapter 5: 95.5%) and only a small degree of variation was observed between the percentages of sequences assigned to a genus (Chapter 3: 79.1%, Chapter 4: 76.2%, Chapter 5: 75.2%). However, sequences in Chapter 4 were around half as likely to be assigned to a species (16.5% in comparison to 36.1% and 32%). Differences in the types of bacteria found in lung samples across results chapters may have had an impact on the taxonomic depth to which sequences could be assigned.

Table 6.2: Comparing error rates, sequences per sample and taxonomic depth across results chapters

	Chapter 3	Chapter 4	Chapter 5
Sequencer used	MiSeq and HiSeq	MiSeq and HiSeq	MiSeq
Sequence error rate	0.39%	0.35%	0.18%
Reduction in sequences due to quality control* (Mean ± SD)	22.4% ± 12.3%	16.2% ± 5.2%	25.8% ± 4.8%
Average sequences per sample (Mean ± SD)	205,625 ± 27,232	138,125 ± 29,306	39,241.5 ± 11,532.2
Good's coverage of samples (Mean ± SD)	0.995 ± 0.00312	0.998 ± 0.000827	0.992 ± 0.00369
Percentage of reads able to be assigned to specific taxonomic depths	Phylum: 99.8% Class: 99.5% Order: 98.5% Family: 95.6% Genus: 79.1% Species: 36.1%	Phylum: 100% Class: 100.0% Order: 99.9% Family: 95.4% Genus: 76.2% Species: 16.5%	Phylum: 99.9% Class: 99.8% Order: 98.6% Family: 95.5% Genus: 75.2% Species: 32.0%

* Includes reductions due to primer removal and quality control carried out within mothur

It is also possible that sequencing errors may have led to errors in taxonomic assignments; the lowest error rate in my results chapters is 0.18% which for a 415 bp sequence would mean that on average there would be 0.747 base errors per sequence. Although it is unlikely that a small amount of sequencing errors would change the taxonomic assignment of sequences at genus level, it is possible that it would affect their assignment at species level as differences between the 16S rRNA gene sequences of species from the same genus are often small or completely absent (134). Also, it is likely

that the diversity of species from less well studied genera represented in 16S databases will be low. For these reasons, species level assignments must be treated with caution; this will be further discussed in the following section.

6.2.2 Differences between mock communities

In Chapter 3 a mock community was sequenced which contained even 16S rRNA gene copies of twenty bacterial species (1:100 dilution). In Chapter 5, a community containing the same species but using staggered quantities of the 16S gene (1,000 to 1,000,000 copies per organism per µl) was sequenced. As the sequencing error rate for these chapters was different, I wanted to check whether this affected the types of biases observed in the mock communities. Bias appears to be consistent between the two communities, with reductions or increases in actual vs expected reads for each species being roughly equivalent (**Table 6.3**). The fact that the majority of the members of the even mock community could be assigned to the correct genus gives me confidence that the protocol I used to generate my results is able to identify a wide range of bacterial taxonomies reliably down to genus level. While the majority of the bacteria in the staggered mock (Chapter 5) were identified correctly to genus level, three were incorrectly identified to species level. In contrast, all of the bacteria identified to species level in the non-staggered community were correctly identified (Chapter 3). This again emphasises the importance of caution when 16S OTUs are assigned to species level.

I was also able to ascertain using the mock community data that, while my Good's coverage values were on average ≥0.992 per chapter, my protocol may still not be identifying rare species. Three of the bacterial species which were present in low numbers in the staggered community were not identified at any taxonomic level: *Actinomyces odontolyticus*, *Bacteroides vulgatus* and *Enterococcus faecalis*. In the even mock community these species were identified down to genus level therefore their absence in the staggered community is likely due to the fact that the sequencing depth was not high enough to capture these low abundance organisms rather than the inability of my protocol to amplify and identify them. It is not possible to assess whether the depth of sequencing in Chapters 3 and 4 would have been enough to identify these low abundance organisms as a staggered community was not included in these chapters. In future studies it might be useful to include both an even and staggered mock community in all sequencing runs in order to assess both bias and the adequacy of sequencing depth.

Table 6.3: Comparing even and staggered mock communities

Species	Even community		Staggered community	
	Abundance (Expected : Actual)	Taxonomy	Abundance (Expected: Actual)	Taxonomy
Acinetobacter baumannii	5% : 1.62%	Genus	0.22% : 0.099%	Misidentified as *A. rhizosphaerae*
Actinomyces odontolyticus	5% : 1.12%	Genus	0.022% : 0%	Not identified
Bacillus cereus	5% : ≤8.60%	Family	2.19% : ≤1.62%	Family
Bacteroides vulgatus	5% : 10.90%	Genus	0.022% : 0%	Not identified
Clostridium beijerinckii	5% : 7.59%	Genus	2.19% : 1.59%	Misidentified as *C.butyricum*
Deinococcus radiodurans	5% : 7.61%	Genus	0.022% : 0.041%	Genus
Enterococcus faecalis	5% : 2.63%	Genus	0.022% : 1.13%	Not identified
Escherichia coli	5% : ≤3.33%	Family	22.01% : ≤38.64%	Family
Helicobacter pylori	5% : 12.65%	Species	0.22% : 4.72%	Species
Lactobacillus gasseri	5% : 6.59%	Genus	0.22% : 0.11%	Genus
Listeria monocytogenes	5% : ≤3.49%	Family	0.22% : ≤0.18%	Family
Neisseria meningitidis	5% : 3.27%	Genus	0.22% : 0.058%	Misidentified as *N.cinerea*
Propionibacterium acnes	5% : 1.77%	Species	0.22% : 0.041%	Species
Pseudomonas aeruginosa	5% : 3.70%	Genus	2.19% : 3.82%	Genus
Rhodobacter sphaeroides	5% : 5.29%	Species	22.01% : 22.00%	Species
Staphylococcus aureus	5% : 1.12%	Species	2.19% : 1.22%	Species
Staphylococcus epidermidis	5% : 1.85%	Species	22.01% : 21.56%	Species
Streptococcus agalactiae	5% : 0.24%	Species	24.12% : 1.38%	Genus
Streptococcus mutans	5% : <1.63%	Genus		
Streptococcus pneumoniae	5% : <1.63%	Genus		
Other	0% : 26.05%	N/A	0% : 0.78%	N/A

6.2.3 Difference between the quantities of DNA in controls and lung samples

Different lung sampling methods may be more or less efficient at sampling the lung microbiota. The quantity of bacterial DNA in the original sample is inversely related to the quantity of contaminating DNA identified in the sample after sequencing (140); therefore, it is important to identify which sampling techniques lead to greater bacterial DNA yield. Samples in all chapters underwent qPCR as described in Section 2.14 to measure the quantity of bacterial DNA present. On average, DNA extraction kit reagent controls across all chapters contained $1.97 \times 10^{-7} \pm 1.10 \times 10^{-7}$ ng/µl while lung samples contained $7.23 \times 10^{-5} \pm 3.89 \times 10^{-4}$ ng/µl. Extraction kit controls contained significantly less bacterial DNA than lung samples (Kruskal-Wallis: P<0.001) (**Fig. 6.1**). Therefore, while contamination from reagents may have contributed towards the observed composition of the bacterial communities in samples, it is likely that the majority of the bacterial DNA in lung samples did not derive from contamination.

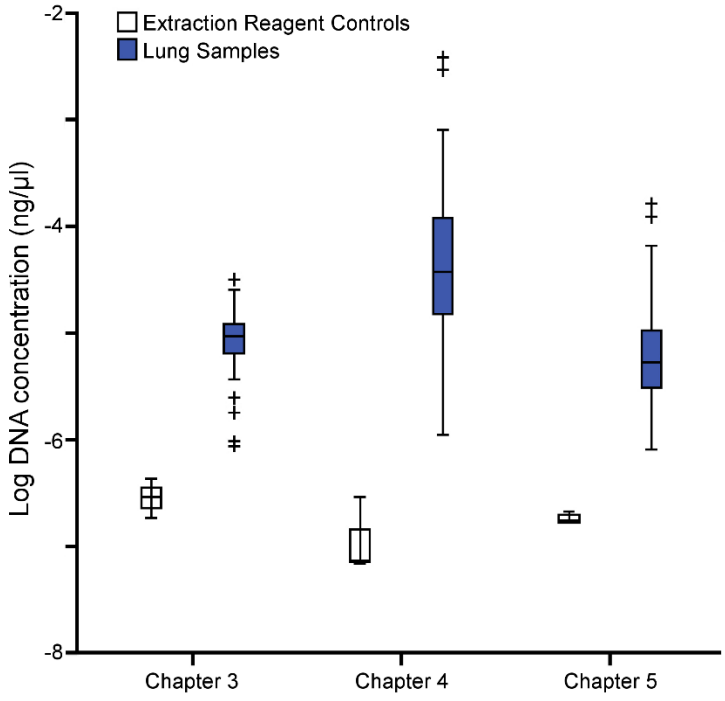


Figure 6.1: Boxplot showing the log concentrations of bacterial DNA found in lung samples and DNA extraction kit reagent controls across results chapters. Amplified DNA consisted of a 179 bp fragment of the V3 region of the 16S rRNA gene. Lung samples from Chapters 3 and 5 were taken via bronchial brushings whereas lung samples in Chapter 4 were taken by post-mortem whole lung PBS washes. Lung samples contained significantly greater quantities of bacterial DNA than reagent controls (Kruskal Wallis: P<0.001) and significantly different concentrations of bacterial DNA across chapters (Kruskal Wallis: P<0.001).

The quantity of bacterial DNA found in lung samples was significantly different between results chapters (Kruskal Wallis: $P<0.001$). The concentration of bacterial DNA measured on average in lung samples was: Chapter 3: $9.90 \times 10^{-6} \pm 5.43 \times 10^{-6}$ ng/µl; Chapter 4: $2.47 \times 10^{-4} \pm 7.48 \times 10^{-4}$ ng/µl and Chapter 5: $1.53 \times 10^{-5} \pm 2.96 \times 10^{-5}$ ng/µl. Lung samples from Chapters 3 and 5 consisted of protected specimen brushings whereas samples from Chapter 4 consisted of whole lung washes which may explain the greater bacterial DNA concentrations found in Chapter 4, as lung washings sample a larger lung area than brushings. Another possible explanation is that there was increased contamination from the PBS used to produce lung washings but this seems unlikely as on average PBS controls contained $1.03 \times 10^{-6} \pm 1.31 \times 10^{-6}$ ng/µl of bacterial DNA, two orders of magnitude lower than the average quantity of DNA found in the lung samples from this chapter.

6.2.4 Identifying 'core' members of the sheep lung microbiota

It has previously been noted that different lots of DNA extraction kit from the same manufacturer can contain different types of bacterial DNA contamination (140). As I used several lots of the Mo Bio Powersoil DNA extraction kit to produce the results contained in my thesis I hypothesised that the types of contaminating DNA found in my reagent controls would change depending on the lot from which they were taken. Unfortunately, extraction kit lot information is not available for all of the DNA extraction kits I used during my thesis. However, it is still possible to make some conclusions based upon the available data.

Differences can be observed in the predominant bacteria found in DNA extraction kit reagent controls based upon the DNA extraction batch in which they were processed (**Fig. 6.2**). Some samples were dominated by specific bacterial genera (*Methylobacterium* for 25/03/2015-1/05/2015 batches or *Burkholderia* for the 28/07/16 batch) whereas others contained more diverse bacterial communities. The controls produced between the 26/07/2016-28/07/2016 all originated from the same extraction kit lot yet differences can still be observed between them.

Salter et al. concluded that the majority of contaminating DNA in extraction kit reagents originated from bacteria which are normally found in water, soil and human skin (140). Genera found at ≥5% abundance in at least one of my DNA extraction kit reagent controls contain some bacterial species which can be isolated from these environments including *Acinetobacter*, *Actinomyces*, *Burkholderia*, *Corynebacterium*, *Enhydrobacter*, *Kocuria*, *Leuconostoc*, *Methylobacterium*, *Micrococcus*, *Paracoccus*, *Pelomonas*, *Propionobacterium*, *Pseudomonas* and *Staphylococcus*. However, my data does not support removing all of the genera/OTUs found in extraction kit controls from my datasets as several of the genera reported in my controls also contain species which can colonise the respiratory tract including *Actinobacillus*, *Aerococcus*, *Aggregatibacter, Anaerococcus*, *Burkholderia*, *Corynebacterium, Gamella*, *Mycobacterium*, *Micrococcus, Pseudomonas* and *Veilonella*. Removing OTUs taxonomically assigned to these genera may mean removing reads belonging to true members

of the lung microbiota. As the majority of my sequencing reads were only assigned to genus and species level assignments are often inaccurate, it is not possible to tell the exact origin of the OTUs found in my controls and there is therefore no way of knowing whether they belonged to species which commonly inhabit environmental or respiratory samples.

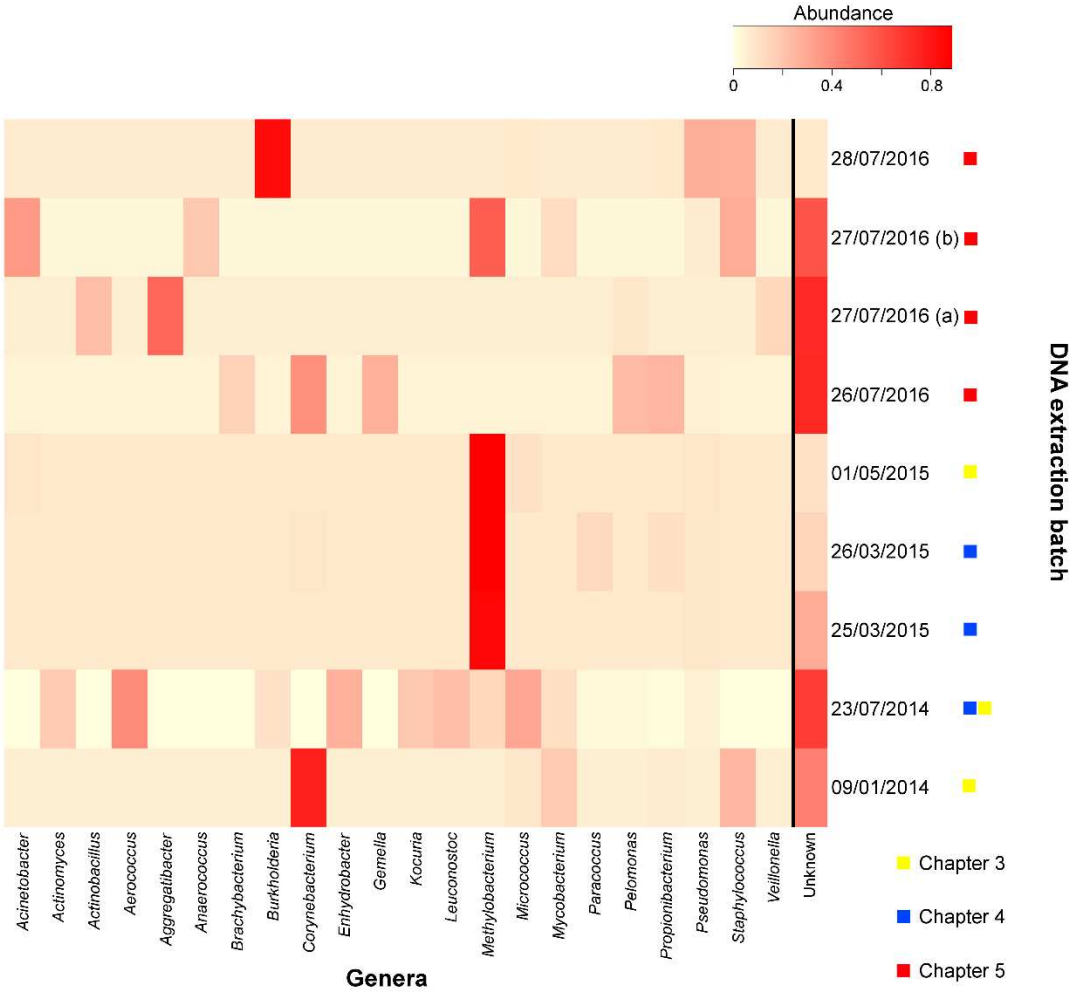


Figure 6.2: Heatmap of bacterial genera present at ≥5% abundance in at least one DNA extraction kit reagent control. The bacterial genera found in reagent controls are highly variable between DNA extraction batches. The chapters to which each control is relevant are indicated by colour. Yellow: Chapter 3, blue: Chapter 4 and red: Chapter 5.

By comparing negative controls and lung samples, it may be possible to identify 'core' lung bacteria which are able to be identified in lung samples across all of my results chapters but are not frequently identified in reagent controls. Several bacterial OTUs occurred on average in >1% of lung samples but in <0.1% of reagent controls in each chapter (**Table 6.4**). The only OTU for which this was the case in all chapters was *Jeotgalicoccus psychrophilus*. It was also found to be the case in two chapters for the

following OTUs: *Staphylococcus sciuri*, *Staphylococcus equorum* and *Mannheimia*. These OTUs were also found in lamb throat swabs produced in Chapter 4: *J.psychrophilus* (0.039% ± 0.10%), *S.sciuri* (0.18% ± 0.56%), *S.equorum* (0.32% ± 1.03%) and *Mannheimia* (8.85% ± 11.19%). It must be emphasised that species level assignments of OTUs using 16S sequences are highly unreliable and we may instead say that it is likely that at least one bacterial species of the genera to which these species belong were common colonisers of the sheep lung in my studies.

Staphylococci and *Mannheimia* are common inhabitants of the upper respiratory tract in farm animals (234). *Jeotgalicoccus* spp. have also previously been found in low abundance in the upper respiratory tracts of sheep, cows, pigs, rabbits, hamsters, cats, humans, dogs and seals (44, 235-237). That this genus has not previously been identified during culture based studies of the sheep lung may be due to the fact that the microbes in the lung are low in abundance or it may be because this genus can easily be mistaken for coagulase-negative *Staphylococcus* spp. using traditional culturing and identification techniques (238). In future studies it might be of interest to attempt to isolate and culture *Jeotgalicoccus* from the sheep lung in order to better characterise it and to analyse whether it plays a role in sheep health.

Table 6.4: OTUs found on average in <0.1% of negative controls but >1% of lung samples

Chapter	OTU	Abundance (%) in lung samples (Mean ± SD)	Abundance (%) in reagent controls (Mean ± SD)
Chapter 3*	*Staphylococcus sciuri*	5.9 ± 5.0	0 ± 0
	Staphylococcus equorum	5.7 ± 6.0	0 ± 0
	Jeotgalicoccus psychrophilus	5.4 ± 6.0	0 ± 0
	Ruminococcaceae	2.6 ± 2.9	0 ± 0
	Granulicatella	1.6 ± 2.5	0 ± 0
	Prevotella	1.6 ± 4.0	0 ± 0
	Bacilli	1.5 ± 4.5	0 ± 0
	Bibersteinia	1.4 ± 5.8	0 ± 0
	Clostridiales	1.1 ± 1.9	0 ±0
	Neisseriaceae	1.1 ± 2.4	0 ± 0
	Staphylococcus epidermidis	1.1 ± 1.9	0 ± 0
Chapter 4	*Staphylococcus equorum*	13.3 ± 9.6	0.01 ± 0.02
	Mannheimia	5.3 ± 21.6	0.03 ± 0.06
	Streptomyces	2.0 ± 3.8	0 ± 0
	Peptostreptococcaceae	1.8 ± 2.2	0.007 ± 0.02
	Jeotgalicoccus psychrophilus	1.6 ± 2.1	0 ± 0
	Microbacterium aurum	1.2 ± 2.8	0 ± 0
	Brevibacterium	1.2 ± 1.4	0 ± 0
	Variovorax paradoxus	1.2 ± 1.8	0.004 ± 0.009
	Turicibacter	1.0 ± 1.8	0 ± 0
Chapter 5	*Mannheimia*	6.5 ± 16.6	0.005 ± 0.006
	Staphylococcus sciuri	5.6 ± 5.1	0.005 ± 0.009
	Pseudomonas veronii	3.9 ± 9.0	0.001 ± 0.003
	Jeotgalicoccus psychrophilus	2.4 ± 3.0	0.002 ± 0.004
	Paracoccus	1.6 ± 7.4	0.004 ± 0.01
	Brachybacterium	1.3 ± 1.6	0.001 ± 0.003

*samples from sheep (n=6) and associated controls from three month longitudinal study only

6.3 Conclusion

When comparing results across my chapters it is important to keep in mind that I identified differences in terms of the error rates, quantity of sequences per sample and taxonomic assignment of sequences between chapters. I also identified significant differences in the quantities of bacterial DNA extracted from lung samples between chapters. Despite these differences, I identified several bacterial OTUs which were abundant in lung samples but not in their corresponding DNA extraction kit reagent controls, namely *J.psychrophilus*, *S.sciuri*, *S.equorum* and *Mannheimia*.

Chapter 7: Discussion

In this thesis I have investigated the composition, diversity and dynamics of the bacterial lung microbiota in sheep. There are still many basic questions left to answer about the bacterial communities in the lungs. For example, whether certain communities are better at providing colonisation resistance against pathogens; whether it is possible to manipulate the lung microbiota through the use of prebiotics or probiotics, and how longitudinally stable the lung microbiota is in the healthy individual. While some of these questions may be able to be answered by studies of humans, others would require the use of disease challenge experiments or the delivery of potentially infectious microorganisms into the lungs. There are obvious ethical problems with performing these types of experiments in humans, therefore a large animal model with a similar respiratory system to humans would be highly useful. The sheep has previously been used as large animal model in various respiratory disease studies due to the similarity of the sheep and human respiratory and immune systems (11, 12), which prompted us to explore whether they could be used to study the lung microbiota.

During my PhD, a greater understanding of the impact which DNA contamination of reagents can have on lung microbiota studies has arisen in the research community. If I were to start my PhD again then I would have included more controls in my study described in Chapter 3, which would have potentially allowed me to explore how the lung microbiota changes over time. However, I did include appropriate negative controls in the studies which followed (Chapter 4 and Chapter 5) and I included mock community controls which allowed me to characterise the inevitable bias which arises from using 16S methodologies to study the microbiota. The presence of bacterial DNA contamination arising from equipment and reagents is unfortunately inevitable in microbiota studies and cannot be removed completely.

Another factor which I would change if I were able to start my PhD again is the variable regions of the 16S gene which were chosen for analysis. At the time our studies were designed the V2-V3 region was thought to be a good choice for studying the lung microbiota as it was thought that the long read produced would allow for our sequences to be identified to a greater taxonomic depth (see Section 2.10). However, the use of an amplicon of this length led to greater sequencing errors due to the small overlap of the forward and reverse reads produced by the Illumina sequencers. It may therefore have been more advantageous to have used a shorter amplicon, such as that which would been produced by amplifying the V4 region of the 16S gene.

Despite these issues, my thesis contains several studies which have allowed us to gain a better understanding of the sheep lung microbiota. In the following sections I will summarise the objectives, results and conclusions of each of my results chapters, along with the limitations associated with each of my studies and possible areas for future research.

7.1 Variability of the Lung Microbiota

In Chapter 3 my objective was to provide one of the first descriptions of the sheep lung microbiota using 16S rRNA gene sequencing and to characterise the inter- and intra-individual variation present. I found that in some sheep there were large differences in the bacterial communities colonising each of the three sampled lung segments, while other sheep contained communities which were similar at all sampling locations (Chapter 3: **Fig. 5**). Although spatial variability was observed, samples still clustered by the sheep from which they were taken indicating that there are host factors which influence lung microbiota composition (Chapter 3: **Fig. 6**). Dickson et al. found similar results in humans and also observed that samples taken from closer to the upper respiratory tract resembled bacterial communities from that niche more than samples which were taken from lower in the respiratory tract (28). In my study, differences were also observed in the lung microbiota based upon sampling depth. When more extensive lung brushing samples were taken from an individual sheep, samples were found to cluster by the depth in the lung from which they were taken. While investigating samples from this sheep I also found that samples which were taken from lung sites located only a few centimetres away from one another did not have the same bacterial community compositions (Chapter 3: **Fig. 7**). This is the first time that the lung microbiota has been examined at this spatial scale in a healthy individual and it leads me to question whether it is possible to routinely include enough sampling locations in lung microbiota studies to take into account the full level of variation which may be present across the lung.

The main limitation inherent to this study is that at the time it was carried out it was not yet well known that different lots of DNA extraction kit contained different types of bacterial DNA (140). As such, a DNA extraction kit control was not included for every lot of kit and for some samples reagents from multiple lots were used. This meant that, while this study included samples taken at three time-points, I could not perform an analysis of the longitudinal variability of the sheep lung microbiota as the samples from each time-point were processed using different lots of kit and at different times. It was therefore not possible to separate out whether the clustering of samples by time-point was due only to differences in the contamination to which the samples had been exposed. Characterising the longitudinal variability of lung bacterial communities is still of interest and could be achieved by repeating this study with appropriate reagent controls and samples which were randomised into DNA extraction batches to prevent false positive results.

7.2 Comparing microbiotas in the upper aerodigestive and lower respiratory tracts of lambs

In Chapter 4 I sampled the lung microbiota in forty lambs in order to gain a better understanding of the types of bacteria which could be found in the sheep lung. I also compared upper and lower respiratory tract samples to see how similar the microbiota at these two sites were to one another. I found that swabs taken from the oropharynx separated into rumen-like and upper respiratory tract-like samples based upon their microbiota compositions. I was unable to conclude whether this represented recent rumination or the leakage of gut contents into the upper respiratory tract post-mortem. This is one of the main limitations of this study as my samples were donated from a separate study where the short time between death and sampling was crucial, and it was of more importance to get the lambs to the dissection table quickly rather than attempting to prevent any leakage of stomach contents while the lambs were being moved.

Regardless, lamb lung fluids did not resemble either the rumen-like or upper respiratory tract-like samples or DNA extraction kit controls in terms of their microbiota composition (**Figs. 4.1** and **4.2**). In humans, although supraglottic samples do cluster separately from lung samples by their microbiota compositions (28), the upper and lower respiratory microbiotas are much more alike than in my samples (20, 23, 49). This may be due to anatomical differences between sheep and humans; for example, the fact that sheep ruminate may mean that they have more anatomical barriers to microaspiration (229). However, my results indicate that microaspiration may still occur in sheep as common members of the ruminal microbiota such as *Prevotella,* Clostridiales, Ruminococcaceae, Lachnospiraceae and *Butyrivibrio* could be identified in lung samples. Other possible differences between sheep and humans which could influence the lung microbiota composition are that sheep produce more saliva, have increased amounts of nasal breathing and have horizontally, rather than vertically, positioned lungs (230, 231).

In future studies it would be interesting to compare samples taken from lambs to those taken from adult sheep from the same flock in order to see if there are any differences in the lung microbiota based on age. A longitudinal study which sampled the lung microbiota of the lamb as it aged in order to better understand how the respiratory microbiota develops would also be of interest and if samples were additionally taken from mother ewes it may be possible to detect if the mother's microbiota plays a role in determining the lung microbiota of the offspring.

7.3 Microbiota in exhaled breath condensate and the lung

In Chapter 5 of my thesis I sought to assess whether exhaled breath condensate samples could be used as a replacement for more invasive lung sampling techniques. I also wanted to discover whether nebulised antibiotic treatment affects the lung microbiota composition. I found that samples produced by bronchial brushings contained significantly more bacterial DNA than samples taken by EBC collection (Chapter 5: **Fig. 2**). This is most likely due to the epithelial lining fluid in EBC being highly diluted with water vapour (239).

As EBC from conscious animals has to pass through the upper respiratory tract but EBC from anaesthetised animals does not, I hypothesised that these samples would cluster separately by their microbiota compositions; however, they did not. This is not to say that these two sample types are interchangeable as EBC samples from the same animal when it was conscious or anaesthetised did not contain the same bacterial communities (Chapter 5: **Fig. 3**). Brushings and EBC samples from anaesthetised animals were found to contain significantly different bacterial communities (Chapter 5: **Fig. 4**), which may have been caused by the fact that EBC samples a larger area of the lung than is sampled by brushings (240). While some studies have found that EBC collection is an efficient method for sampling lung microbes (97, 241) other studies have found that it is inefficient at detecting the presence of potentially pathogenic lung microorganisms (242-245). Although my results lead me to conclude that EBC cannot be used to replace more invasive lung sampling methods when studying the lung microbiota, there are other potential uses for EBC collection in 16S studies. For example, collecting EBC from animals which are at risk of bovine respiratory disease or pasteurellosis could be useful for monitoring the types of exhaled bacteria which these animals are exposing each other to.

During this study we also treated the sheep with colistimethate sodium and sampled the respiratory tract pre and post treatment. Although relatively low concentrations of colistin were detected in the epithelial lining fluid, brushing samples were found to cluster separately pre and post treatment by their microbiota compositions (Chapter 5: **Fig. 5**). Previously we showed that injection with colistimethate sodium was correlated with changes in the sheep lung microbiota (115). Studies examining the effects of antibiotic treatment on the human lung microbiota have all been carried out in patients with lung disease, therefore it is difficult to compare them directly to our findings. In two studies looking at changes in the lung flora due to antibiotic treatment, one in cystic fibrosis patients and one in cattle affected by bovine respiratory disease (74, 132), it was found that the effects of antibiotic treatment were transient. Both of these studies were carried out in individuals whose lungs were infected by pathogenic microorganisms and where antibiotic treatment was designed specifically to reduce the number of these organisms. To discover whether antibiotic effects are transient in the healthy lung, a similar experiment to the one described in this results chapter could be carried out, with a larger amount of sampling points post treatment. This would enable us to assess whether it is possible to cause a long-term change in the healthy lung microbiota.

7.4 Identifying core members of the sheep lung microbiota

In my final results chapter I analysed the inter-study variability across my three previous results chapters and attempted to identify potential 'core' members of the sheep lung microbiota. I found that there were differences between my results chapters in terms of the sequencing error rates, taxonomic identifications of sequences, coverage and sequence numbers per sample (**Table 6.2**). Some of this variation is likely to be due to the fact that different sequencing platforms, numbers of sequencing runs and versions of Illumina chemistry were used in the three chapters. I do not feel confident in stating that the species level assignments in my chapters are accurate as the sequencing error rates indicate that it is likely that errors will have been present in many of my sequences. As species within the same genus can have highly similar 16S rRNA gene sequences it is logical to treat species level assignments with caution. I have more confidence that my protocol is able to correctly assign sequences to genus as the mock community genus level assignments were all correct.

Consistent biases were identified in my mock communities (**Table 6.3**) which may have been caused by my DNA extraction, PCR amplification, sequencing or bioinformatic methodologies (or by a combination of these) (158). One identified bias which is of some concern is the underrepresentation of *Streptococcus* species. *Streptococcus* are common members of the upper respiratory tract microbiota in many animals and it is therefore plausible that they might also form part of the lung microbiota. It is also possible that some rarer members of the lung microbiota will not have been identified in my samples as several of the rarer members of my staggered mock community were not found in my dataset after sequencing, despite the apparently high coverage values obtained for all of my samples. I would therefore recommend including in every sequencing run two mock communities, one with equal copies of the 16S genes from each bacterial species and one containing staggered quantities of the 16S genes from the same bacterial species. This will allow for both the identification of biases and for the level of coverage to be better characterised.

DNA contamination originating from reagents or lab equipment may also cause bias in studies of low biomass microbial communities. While attempts were made to avoid the contamination of my samples some degree of contamination was practically unavoidable (**Fig. 6.2**). Salter et al. found that the majority of bacterial species in DNA extraction kits originated from skin, soil and water environments (140). Several of the bacteria I identified in my reagent only controls may have arisen from these environments; however, there were also many genera present which contain species which have been shown to colonise the respiratory tract. Where reagent only controls were dominated by one bacterial OTU which clearly would not colonise the lung (eg. *Methylobacterium komagatae* in Chapter 3) I felt confident in removing these OTUs from my datasets. However, if reagent controls contained a diversity of bacterial OTUs, which may or may not be colonisers of the respiratory tract, I did not

remove these OTUs from my datasets as I felt that this may lead to the removal of true members of the lung microbiota. While methods have been developed which attempt to remove contaminating DNA from datasets using sequence abundance data (39, 154), I did not use these methods as they have not yet been extensively adopted and tested. However, as future work it may be interesting to explore my datasets using these methods.

Despite methodological variability between results chapters and the presence of bias and contamination, I was able to identify four OTUs which were common in lung samples but not in DNA extraction controls: *J.psychrophilus*, *S.sciuri*, *S.equorum* and *Mannheimia*. As previously stated, these species level assignments may not be accurate but we may be more confident that the genus level assignments are correct. Staphylococci and *Mannheimia* are both common members of the upper respiratory tract microbiota of farm animals (234). *Jeotgalicoccus* is a poorly studied genus (238) which has previously been identified in the upper respiratory tracts of several mammals including farm animals (44, 235-237). Using culturing and classical identification techniques these bacteria are indistinguishable from coagulase-negative *Staphylococcus* spp. (238) which may explain the fact that they have not previously been identified in the sheep lung during culture based studies. In future studies the isolation and characterisation of this *Jeotgalicoccus* species may be of interest, in order to establish whether it plays a role in host health and to assess whether the species level assignment of *J.psychrophilus* is accurate. More accurate species level assignments could also be achieved by sequencing the full 16S rRNA genes of bacteria from a wider variety of environments and by building 16S reference databases which contain bacterial species specific to particular environmental niches (246).

7.5 General conclusions

My thesis contains the first extensive description of the sheep lung microbiota as characterised by 16S rRNA gene sequencing. While I found that there are some differences between the sheep and human lung microbiotas I feel that the sheep does have a use as a large animal model for studying the lung microbiota, including the assessment of different sampling techniques, as performed in Chapter 5.

While my thesis contributes to our understanding of the healthy sheep lung microbiota, the scope of my thesis does not allow me to make conclusions as to whether the lung microbiota has an impact on the health of sheep or whether there are changes in the sheep lung microbiota during respiratory disease. This would be an interesting avenue for future research, as would studies in other livestock species.

I feel that the eventual goal of lung microbiota research should be to provide treatments for diseases which involve lung microbial communities. These therapies may consist of probiotics, aerosolised

antibiotics, antibiotic treatments targeted at specific members of the lung microbiota or bacterial replacement therapy, such as is performed on the gut microbiota during faecal transplantation. However, in comparison to other areas of microbiota research the lung microbiota field is still in its infancy and it is likely to be some time before the complex interactions between the host, their environment and lung microbial communities are sufficiently understood to allow for the development of these kind of therapies.

Chapter 8: Bibliography

1. **Glendinning L, Wright S, Pollock J, Tennant P, Collie D, McLachlan G.** 2016. Variability of the sheep lung microbiota. Appl Environ Microbiol **82:**3225-3238.
2. **Glendinning L, Wright S, Tennant P, Gill AC, Collie D, McLachlan G.** 2017. Microbiota in exhaled breath condensate and the lung. Appl Environ Microbiol **83:**e00515-00517.
3. **Darwin C.** 1868. The Variation of Animals and Plants Under Domestication: Volume 2. Organe Judd, New York, US.
4. **Cabreiro F, Gems D.** 2013. Worms need microbes too: Microbiota, health and aging in *Caenorhabditis elegans*. EMBO Mol Med **5:**1300-1310.
5. **Hacquard S, Garrido-Oter R, Gonzalez A, Spaepen S, Ackermann G, Lebeis S, McHardy AC, Dangl JL, Knight R, Ley R, Schulze-Lefert P.** 2015. Microbiota and host nutrition across plant and animal kingdoms. Cell Host Microbe **17:**603-616.
6. **Lee YK, Mazmanian SK.** 2010. Has the microbiota played a critical role in the evolution of the adaptive immune system? Science **330:**1768-1773.
7. **Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI.** 2008. Worlds within worlds: Evolution of the vertebrate gut microbiota. Nat Rev Micro **6:**776-788.
8. **Man WH, de Steenhuijsen Piters WAA, Bogaert D.** 2017. The microbiota of the respiratory tract: Gatekeeper to respiratory health. Nat Rev Micro **15:**259-270.
9. **McKenney Peter T, Pamer Eric G.** 2015. From hype to hope: The gut microbiota in enteric infectious disease. Cell **163:**1326-1332.
10. **Tlaskalova-Hogenova H, Stepankova R, Kozakova H, Hudcovic T, Vannucci L, Tuckova L, Rossmann P, Hrncir T, Kverka M, Zakostelska Z, Klimesova K, Pribylova J, Bartova J, Sanchez D, Fundova P, Borovska D, Srutkova D, Zidek Z, Schwarzer M, Drastich P, Funda DP.** 2011. The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: Contribution of germ-free and gnotobiotic animal models of human diseases. Cell Mol Immunol **8:**110-120.
11. **Scheerlinck JPY, Snibson KJ, Bowles VM, Sutton P.** 2008. Biomedical applications of sheep models: From asthma to vaccines. Trends Biotechnol **26:**259-266.
12. **Meeusen EN, Snibson KJ, Hirst SJ, Bischof RJ.** 2009. Sheep as a model species for the study and treatment of human asthma and other respiratory diseases. Drug Discov Today Dis Models **6:**101-106.
13. **O'Dwyer DN, Dickson RP, Moore BB.** 2016. The lung microbiome, immunity, and the pathogenesis of chronic lung disease. J Immunol **196:**4839-4847.
14. **Dickson RP, Huffnagle GB.** 2015. The lung microbiome: New principles for respiratory bacteriology in health and disease. PLoS Pathog **11:**e1004923.
15. **Meyer KC.** 2002. Lung Immunology and Host Defense, p 332-345. *In* Bittar EE (ed), Pulmonary Biology in Health and Disease. Springer, New York, US.
16. **Burnett D.** 1986. Immunoglobulins in the lung. Thorax **41:**337-344.

17. **Lecaille F, Lalmanach G, Andrault PM.** 2016. Antimicrobial proteins and peptides in human lung diseases: A friend and foe partnership with host proteases. Biochimie **122:**151-168.
18. **Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, Davies J, Ervine A, Poulter L, Pachter L, Moffatt MF, Cookson WOC.** 2010. Disordered microbial communities in asthmatic airways. PLoS ONE **5:**e8578.
19. **Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, Young VB, Toews GB, Curtis JL, Sundaram B, Martinez FJ, Huffnagle GB.** 2011. Analysis of the lung microbiome in the "healthy" smoker and in COPD. PLoS ONE **6:**e16384.
20. **Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG.** 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med **184:**957-963.
21. **Dickson RP, Erb-Downward JR, Huffnagle GB.** 2015. Homeostasis and its disruption in the lung microbiome. Am J Physiol Lung Cell Mol Physiol **309:**L1047-L1055.
22. **Marsland BJ, Gollwitzer ES.** 2014. Host-microorganism interactions in lung diseases. Nat Rev Immunol **14:**827-835.
23. **Venkataraman A, Bassis CM, Beck JM, Young VB, Curtis JL, Huffnagle GB, Schmidt TM.** 2015. Application of a neutral community model to assess structuring of the human lung microbiome. mBio **6:**e0084-0014.
24. **Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM.** 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med **187:**1067-1075.
25. **Cui L, Lucht L, Tipton L, Rogers MB, Fitch A, Kessinger C, Camp D, Kingsley L, Leo N, Greenblatt RM, Fong S, Stone S, Dermand J, Kleerup E, Huang L, Morris A, Ghedin E.** 2015. Topographical diversity of the respiratory tract mycobiome and alteration in HIV and lung disease. Am J Respir Crit Care Med **191:**932-942.
26. **MacArthur RH, Wilson EO.** 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, US.
27. **Dickson RP, Erb-Downward JR, Huffnagle GB.** 2014. Towards an ecology of the lung: New conceptual models of pulmonary microbiology and pneumonia pathogenesis. Lancet Resp Med **2:**238-246.
28. **Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Beck JM, Huffnagle GB, Curtis JL.** 2015. Spatial variation in the healthy human lung microbiome and the adapted island model of lung biogeography. Ann Am Thorac Soc **12:**821-830.

29. **Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, Curtis JL.** 2017. Bacterial topography of the healthy human lower respiratory tract. mBio **8:**e02287-02216.
30. **Dickson RP, Singer BH, Newstead MW, Falkowski NR, Erb-Downward JR, Standiford TJ, Huffnagle GB.** 2016. Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. Nat Microbiol **1:**16113.
31. **Gollwitzer ES, Saglani S, Trompette A, Yadava K, Sherburn R, McCoy KD, Nicod LP, Lloyd CM, Marsland BJ.** 2014. Lung microbiota promotes tolerance to allergens in neonates via PD-L1. Nat Med **20:**642-647.
32. **Karlstrom A, Lindgren H, Hildingsson I.** 2013. Maternal and infant outcome after caesarean section without recorded medical indication: Findings from a Swedish casecontrol study. BJOG **120:**479-486.
33. **Thavagnanam S, Fleming J, Bromley A, Shields MD, Cardwell CR.** 2008. A meta-analysis of the association between Caesarean section and childhood asthma. Clin Exp Allergy **38:**629-633.
34. **Bosch AATM, Levin E, van Houten MA, Hasrat R, Kalkman G, Biesbroek G, de Steenhuijsen Piters WAA, de Groot P-KCM, Pernet P, Keijser BJF, Sanders EAM, Bogaert D.** 2016. Development of upper respiratory tract microbiota in infancy is affected by mode of delivery. EBioMedicine **9:**336-345.
35. **Holgerson PL, Harnevik L, Hernell O, Tanner ACR, Johansson I.** 2011. Mode of birth delivery affects oral microbiota in infants. J Dent Res **90:**1183-1188.
36. **Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R.** 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci USA **107:**11971-11975.
37. **Pezzulo AA, Kelly PH, Nassar BS, Rutland CJ, Gansemer ND, Dohrn CL, Costello AJ, Stoltz DA, Zabner J.** 2013. Abundant DNase I-sensitive bacterial DNA in healthy porcine lungs and Its implications for the lung microbiome. Appl Environ Microbiol **79:**5936-5941.
38. **Rogers GB, Cuthbertson L, Hoffman LR, Wing PA, Pope C, Hooftman DA, Lilley AK, Oliver A, Carroll MP, Bruce KD.** 2013. Reducing bias in bacterial community analysis of lower respiratory infections. ISME J **7:**697-706.
39. **Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB.** 2015. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. mBio **6:**e00037-00015.
40. **Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu W-H, Lakshmanan A, Wade WG.** 2010. The human oral microbiome. J Bacteriol **192:**5002-5017.
41. **Wade WG.** 2013. The oral microbiome in health and disease. Pharmacol Res **69:**137-143.

42. **Dewhirst FE, Klein EA, Thompson EC, Blanton JM, Chen T, Milella L, Buckley CMF, Davis IJ, Bennett M-L, Marshall-Jones ZV.** 2012. The canine oral microbiome. PLoS ONE **7:**e36067.

43. **Sturgeon A, Pinder SL, Costa MC, Weese JS.** 2014. Characterization of the oral microbiota of healthy cats using next-generation sequencing. Vet J **201:**223-229.

44. **Kittelmann S, Kirk MR, Jonker A, McCulloch A, Janssen PH.** 2015. Buccal swabbing as a noninvasive method to determine bacterial, archaeal, and eukaryotic microbial community structures in the rumen. Appl Environ Microbiol **81:**7470-7483.

45. **Chun J, Kim KY, Lee JH, Choi Y.** 2010. The analysis of oral microbial communities of wild-type and toll-like receptor 2-deficient mice using a 454 GS FLX Titanium pyrosequencer. BMC Microbiol **10:**101.

46. **Hyde ER, Luk B, Cron S, Kusic L, McCue T, Bauch T, Kaplan H, Tribble G, Petrosino JF, Bryan NS.** 2014. Characterization of the rat oral microbiome and the effects of dietary nitrate. Free Radic Biol Med **77:**249-257.

47. **Stahringer SS, Clemente JC, Corley RP, Hewitt J, Knights D, Walters WA, Knight R, Krauter KS.** 2012. Nurture trumps nature in a longitudinal survey of salivary bacterial communities in twins from early adolescence to early adulthood. Genome Res **22:**2146-2152.

48. **Hajishengallis G.** 2015. Periodontitis: From microbial immune subversion to systemic inflammation. Nat Rev Immunol **15:**30-44.

49. **Segal L, Alekseyenko A, Clemente J, Kulkarni R, Wu B, Chen H, Berger K, Goldring R, Rom W, Blaser M, Weiden M.** 2013. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. Microbiome **1:**19.

50. **Segal LN, Clemente JC, Tsay J-CJ, Koralov SB, Keller BC, Wu BG, Li Y, Shen N, Ghedin E, Morris A, Diaz P, Huang L, Wikoff WR, Ubeda C, Artacho A, Rom WN, Sterman DH, Collman RG, Blaser MJ, Weiden MD.** 2016. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. Nat Microbiol **1:**16031.

51. **Yun Y, Srinivas G, Kuenzel S, Linnenbrink M, Alnahas S, Bruce KD, Steinhoff U, Baines JF, Schaible UE.** 2014. Environmentally determined differences in the murine lung microbiota and their relation to alveolar architecture. PLoS ONE **9:**e113466.

52. **Dickson R, Erb-Downward J, Prescott H, Martinez F, Curtis J, Lama V, Huffnagle G.** 2014. Cell-associated bacteria in the human lung microbiome. Microbiome **2:**28.

53. **Budden KF, Gellatly SL, Wood DLA, Cooper MA, Morrison M, Hugenholtz P, Hansbro PM.** 2017. Emerging pathogenic links between microbiota and the gut-lung axis. Nat Rev Micro **15:**55–63.

54. **Samuelson DR, Welsh DA, Shellito JE.** 2015. Regulation of lung immunity and host defense by the intestinal microbiota. Front Microbiol **6:**1085.

55. **Dickson RP, Cox MJ.** 2016. Gut microbiota and protection from pneumococcal pneumonia. Gut **66:**384.

56. **Rylance J, Kankwatira A, Nelson DE, Toh E, Day RB, Lin HY, Gao X, Dong QF, Sodergren E, Weinstock GM, Heyderman RS, Twigg HL, Gordon SB.** 2016. Household air pollution and the lung microbiome of healthy adults in Malawi: A cross-sectional study. BMC Microbiol **16:**182.

57. **Heijerman H.** 2005. Infection and inflammation in cystic fibrosis: A short review. J Cyst Fibros **4, Supplement 2:**3-5.

58. **Madan JC, Koestler DC, Stanton BA, Davidson L, Moulton LA, Housman ML, Moore JH, Guill MF, Morrison HG, Sogin ML, Hampton TH, Karagas MR, Palumbo PE, Foster JA, Hibberd PL, O'Toole GA.** 2012. Serial analysis of the gut and respiratory microbiome in cystic fibrosis in infancy: Interaction between intestinal and respiratory tracts and impact of nutritional exposures. mBio **3:**e00251-00212.

59. **Frayman KB, Armstrong DS, Carzino R, Ferkol TW, Grimwood K, Storch GA, Teo SM, Wylie KM, Ranganathan SC.** 2017. The lower airway microbiota in early cystic fibrosis lung disease: A longitudinal analysis. **EPub ahead of print:**doi:10.1136/thoraxjnl-2016-209279.

60. **Coburn B, Wang PW, Diaz Caballero J, Clark ST, Brahma V, Donaldson S.** 2015. Lung microbiota across age and disease stage in cystic fibrosis. Sci Rep **5:**10241.

61. **Lyczak JB, Cannon CL, Pier GB.** 2002. Lung infections associated with cystic fibrosis. Clin Microbiol Rev **15:**194-222.

62. **Pernet E, Guillemot L, Burgel P-R, Martin C, Lambeau G, Sermet-Gaudelus I, Sands D, Leduc D, Morand PC, Jeammet L, Chignard M, Wu Y, Touqui L.** 2014. *Pseudomonas aeruginosa* eradicates *Staphylococcus aureus* by manipulating the host immunity. Nat Commun **5:**5105.

63. **Filkins LM, Graber JA, Olson DG, Dolben EL, Lynd LR, Bhuju S, O'Toole GA.** 2015. Coculture of *Staphylococcus aureus* with *Pseudomonas aeruginosa* drives *S.aureus* towards fermentative metabolism and reduced viability in a cystic fibrosis model. J Bacteriol **197:**2252-2264.

64. **Bacci G, Paganin P, Lopez L, Vanni C, Dalmastri C, Cantale C, Daddiego L, Perrotta G, Dolce D, Morelli P, Tuccio V, De Alessandri A, Fiscarelli EV, Taccetti G, Lucidi V, Bevivino A, Mengoni A.** 2016. Pyrosequencing unveils cystic fibrosis lung microbiome differences associated with a severe lung function decline. PLoS ONE **11:**e0156807.

65. **Garcia JRC, Krause A, Schulz S, Rodriguez-Jimenez FJ, Kluver E, Adermann K, Forssmann U, Frimpong-Boateng A, Bals R, Forssmann WG.** 2001. Human beta-defensin 4: A novel inducible peptide with a specific salt-sensitive spectrum of antimicrobial activity. FASEB J **15:**1819-1891.

66. **Goldman MJ, Anderson GM, Stolzenberg ED, Kari UP, Zasloff M, Wilson JM.** 1997. Human β-defensin-1 Is a salt-sensitive antibiotic in lung that is inactivated in cystic fibrosis. Cell **88:**553-560.

67. **Bals R, Wang XR, Wu ZR, Freeman T, Bafna V, Zasloff M, Wilson JM.** 1998. Human beta-defensin 2 is a salt-sensitive peptide antibiotic expressed in human lung. J Clin Invest **102:**874-880.

68. **Stoltz DA, Meyerholz DK, Welsh MJ.** 2015. Origins of cystic fibrosis lung disease. N Engl J Med **372:**351-362.

69. **Starner TD, Zhang N, Kim GH, Apicella MA, McCray PB.** 2006. *Haemophilus influenzae* forms biofilms on airway epithelia - Implications in cystic fibrosis. Am J Respir Crit Care Med **174:**213-220.

70. **Bjarnsholt T, Jensen PØ, Fiandaca MJ, Pedersen J, Hansen CR, Andersen CB, Pressler T, Givskov M, Høiby N.** 2009. *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. Pediatr Pulmonol **44:**547-558.

71. **Matsui H, Wagner VE, Hill DB, Schwab UE, Rogers TD, Button B, Taylor RM, Superfine R, Rubinstein M, Iglewski BH, Boucher RC.** 2006. A physical linkage between cystic fibrosis airway surface dehydration and *Pseudomonas aeruginosa* biofilms. Proc Natl Acad Sci USA **103:**18131-18136.

72. **Fodor AA, Klem ER, Gilpin DF, Elborn JS, Boucher RC, Tunney MM, Wolfgang MC.** 2012. The adult cystic fibrosis airway microbiota is stable over time and infection type, and highly resilient to antibiotic treatment of exacerbations. PLoS ONE **7:**e45001.

73. **Cuthbertson L, Rogers GB, Walker AW, Oliver A, Green LE, Daniels TWV, Carroll MP, Parkhill J, Bruce KD, van der Gast CJ.** 2016. Respiratory microbiota resistance and resilience to pulmonary exacerbation and subsequent antimicrobial intervention. ISME J **10:**1081-1091.

74. **Li J, Hao CY, Ren LL, Xiao Y, Wang JW, Qin XM.** 2016. Data mining of lung microbiota in cystic fibrosis patients. PLoS ONE **11:**e0164510.

75. **Whelan FJ, Heirali AA, Rossi L, Rabin HR, Parkins MD, Surette MG.** 2017. Longitudinal sampling of the lung microbiota in individuals with cystic fibrosis. PLoS ONE **12:**e0172811.

76. **Jorth P, Staudinger Benjamin J, Wu X, Hisert KB, Hayden H, Garudathri J, Harding Christopher L, Radey Matthew C, Rezayat A, Bautista G, Berrington William R, Goddard Amanda F, Zheng C, Angermeyer A, Brittnacher Mitchell J, Kitzman J, Shendure J, Fligner Corinne L, Mittler J, Aitken Moira L, Manoil C, Bruce James E, Yahr Timothy L, Singh Pradeep K.** 2015. Regional isolation drives bacterial diversification within cystic fibrosis lungs. Cell Host Microbe **18:**307-319.

77. **Willner D, Haynes MR, Furlan M, Schmieder R, Lim YW, Rainey PB, Rohwer F, Conrad D.** 2012. Spatial distribution of microbial communities in the cystic fibrosis lung. ISME J **6:**471-474.

78. **Zemanick ET, Wagner BD, Robertson CE, Stevens MJ, Szefler SJ, Accurso FJ, Sagel SD, Harris JK.** 2015. Assessment of airway microbiota and inflammation in cystic fibrosis using multiple sampling methods. Ann Am Thorac Soc **12:**221-229.

79. **Hogan DA, Willger SD, Dolben EL, Hampton TH, Stanton BA, Morrison HG, Sogin ML, Czum J, Ashare A.** 2016. Analysis of lung microbiota in bronchoalveolar lavage, protected brush and sputum samples from subjects with mild-to-moderate cystic fibrosis lung disease. PLoS ONE **11:**e0149998.

80. **Rogers DF.** 2005. Mucociliary dysfunction in COPD: Effect of current pharmacotherapeutic options. Pulm Pharmacol Ther **18:**1-8.

81. **Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE.** 2012. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. PLoS ONE **7:**e47305.

82. **Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, Cooper J, Sin DD, Mohn WW, Hogg JC.** 2012. The lung tissue microbiome in chronic obstructive pulmonary disease. Am J Respir Crit Care Med **185:**1073-1080.

83. **Einarsson GG, Comer DM, McIlreavey L, Parkhill J, Ennis M, Tunney MM, Elborn JS.** 2016. Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. Thorax **71:**795-803.

84. **Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SAG, Homola D, Trujillo-Torralbo M-B, Elkin S, Kon OM, Cookson WOC, Moffatt MF, Johnston SL.** 2013. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. Am J Respir Crit Care Med **188:**1224-1231.

85. **Richmond BW, Brucker RM, Han W, Du R-H, Zhang Y, Cheng D-S, Gleaves L, Abdolrasulnia R, Polosukhina D, Clark PE, Bordenstein SR, Blackwell TS, Polosukhin VV.** 2016. Airway bacteria drive a progressive COPD-like phenotype in mice with polymeric immunoglobulin receptor deficiency. Nat Commun **7:**11240.

86. **Syed SA, Whelan FJ, Waddell B, Rabin HR, Parkins MD, Surette MG.** 2016. Re-emergence of lower airway microbiota in cystic fibrosis lung transplant patients. Ann Am Thorac Soc **13:**2132-2142.

87. **Bernasconi E, Pattaroni C, Koutsokera A, Pison C, Kessler R, Benden C, Soccal PM, Magnan A, Aubert J-D, Marsland BJ, Nicod LP.** 2016. Airway microbiota determines innate cell inflammatory or tissue remodeling profiles in lung transplantation. Am J Respir Crit Care Med **194:**1252-1263.

88. **Hooper LV, Littman DR, Macpherson AJ.** 2012. Interactions between the microbiota and the immune system. Science **336:**1268-1273.

89. **Kamada N, Seo SU, Chen GY, Nunez G.** 2013. Role of the gut microbiota in immunity and inflammatory disease. Nat Rev Immunol **13:**321-335.

90. **Macpherson AJ, Gatto D, Sainsbury E, Harriman GR, Hengartner H, Zinkernagel RM.** 2000. A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria. Science **288:**2222-+.

91. **Powrie F, Leach MW, Mauze S, Caddle LB, Coffman RL.** 1993. Phenotypically distinct subsets of CD4(+) T-cells induce or protect from chronic intestinal inflammation in C - B-17 SCID mice. Int Immunol **5:**1461-1471.

92. **Salzman NH, Ghosh D, Huttner KM, Paterson Y, Bevins CL.** 2003. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. Nature **422:**522-526.

93. **Pabst O, Herbrand H, Friedrichsen M, Velaga S, Dorsch M, Berhardt G, Worbs T, Macpherson AJ, Forster R.** 2006. Adaptation of solitary intestinal lymphoid tissue in response to microbiota and chemokine receptor CCR7 signaling. J Immunol **177:**6824-6832.

94. **Ivanov, II, Atarashi K, Manel N, Brodie EL, Shima T, Karaoz U, Wei DG, Goldfarb KC, Santee CA, Lynch SV, Tanoue T, Imaoka A, Itoh K, Takeda K, Umesaki Y, Honda K, Littman DR.** 2009. Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell **139:**485-498.

95. **Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, Iwasaki A.** 2011. Microbiota regulates immune defense against respiratory tract influenza A virus infection. **108**.

96. **Twigg HL, Knox KS, Zhou J, Crothers KA, Nelson DE, Toh E, Day RB, Lin HY, Gao X, Dong QF, Mi DM, Katz BP, Sodergren E, Weinstock GM.** 2016. Effect of advanced HIV infection on the respiratory microbiome. Am J Respir Crit Care Med **194:**226-235.

97. **Carpagnano GE, Malerba M, Lacedonia D, Susca A, Logrieco A, Carone M, Cotugno G, Palmiotti GA, Foschino-Barbaro MP.** 2016. Analysis of the fungal microbiome in exhaled breath condensate of patients with asthma. Allergy Asthma Proc **37:**E41-E46.

98. **Poroyko V, Meng F, Meliton A, Afonyushkin T, Ulanov A, Semenyuk E, Latif O, Tesic V, Birukova AA, Birukov KG.** 2015. Alterations of lung microbiota in a mouse model of LPS-induced lung injury. Am J Physiol Lung Cell Mol Physiol **309:**L76-L83.

99. **Boyton RJ, Reynolds CJ, Quigley KJ, Altmann DM.** 2013. Immune mechanisms and the impact of the disrupted lung microbiome in chronic bacterial lung infection and bronchiectasis. Clin Exp Immunol **171:**117-123.

100. **Kelly BJ, Imai I, Bittinger K, Laughlin A, Fuchs BD, Bushman FD, Collman RG.** 2016. Composition and dynamics of the respiratory tract microbiome in intubated patients. Microbiome **4:**7.

101. **Zakharkina T, Martin-Loeches I, Matamoros S, Povoa P, Torres A, Kastelijn JB, Hofstra JJ, de Wever B, de Jong M, Schultz MJ, Sterk PJ, Artigas A, Bos LDJ.** 2017. The dynamics of the pulmonary microbiome during mechanical ventilation in the intensive care unit and the association with occurrence of pneumonia. Thorax **Epub ahead of print:**doi:10.1136/thoraxjnl-2016-209158.

102. **Taylor SL, Woodman RJ, Chen AC, Burr LD, Gordon DL, McGuckin MA, Wesselingh S, Rogers GB.** 2017. FUT2 genotype influences lung function, exacerbation frequency and airway microbiota in non-CF bronchiectasis. Thorax **72:**304-310.

103. **Audfray A, Varrot A, Imberty A.** 2013. Bacteria love our sugars: Interaction between soluble lectins and human fucosylated glycans, structures, thermodynamics and design of competing glycocompounds. C R Chim **16:**482-490.

104. **Benson AK, Kelly SA, Legge R, Ma FR, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua KJ, Kachman SD, Moriyama EN, Walter J, Peterson DA, Pomp D.** 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A **107:**18933-18938.

105. **Rogers GB, Zain NMM, Bruce KD, Burr LD, Chen AC, Rivett DW, McGuckin MA, Serisier DJ.** 2014. A novel microbiota stratification system predicts future exacerbations in bronchiectasis. Ann Am Thorac Soc **11:**496-503.

106. **Rogers GB, van der Gast CJ, Serisier DJ.** 2015. Predominant pathogen competition and core microbiota divergence in chronic airway infection. ISME J **9:**217-225.

107. **Durack J, Lynch SV, Nariya S, Bhakta NR, Beigelman A, Castro M, Dyer A-M, Israel E, Kraft M, Martin RJ, Mauger DT, Rosenberg SR, Sharp-King T, White SR, Woodruff PG, Avila PC, Denlinger LC, Holguin F, Lazarus SC, Lugogo N, Moore WC, Peters SP, Que L, Smith LJ, Sorkness CA, Wechsler M, Wenzel SE, Boushey HA, Huang YJ.** 2016. Features of the bronchial bacterial microbiome associated with atopy, asthma and responsiveness to inhaled corticosteroid treatment. J Allergy Clin Immunol **EPub ahead of print:**doi: 10.1016/j.jaci.2016.1008.1055.

108. **Van der Velden J, Snibson KJ.** 2011. Airway disease: The use of large animal models for drug discovery. Pulm Pharmacol Ther **24:**525-532.

109. **Davies P, Reid L, Lister G, Pitt B.** 1988. Postnatal growth of the sheep lung: A morphometric study. Anat Rec **220:**281-286.

110. **Van der Velden J, Harkness LM, Barker DM, Barcham GJ, Ugalde CL, Koumoundouros E, Bao H, Organ LA, Tokanovic A, Burgess JK, Snibson KJ.** 2016. The effects of tumstatin on vascularity, airway inflammation and lung function in an experimental sheep model of chronic asthma. Sci Rep **6:**26309.

111. **Roche E, Passmore M, Simonova G, Diab S, Dunster K, Bierman W, Fraser J, Tung J-P.** 2016. A histologic approach to qualify lung tissue damage in a sheep model of transfusion-related lung injury: Role of red blood cell storage duration and heat treatment. Am J Clin Pathol **146: Supplement 1:**S20-S20.

112. **Velasquez DRB, Teixeira-Neto FJ, Lagos-Carvajal AP, Steim-Diniz M, Rodriguez NC, Dias CA.** 2015. Effects of different inspired oxygen fractions on sildenafil-induced pulmonary anti-hypertensive effects in a sheep model of acute pulmonary embolism. Life Sci **133:**53-53.

113. **Liu SQ, Yi Y, Wang MH, Chen QH, Huang YZ, Liu L, Xie JF, Zhou DY, Qiu HB.** 2013. Higher frequency ventilation attenuates lung Injury during high-frequency oscillatory ventilation in sheep models of acute respiratory distress syndrome. Anesthesiology **119:**398-411.

114. **McLachlan G, Davidson H, Holder E, Davies LA, Pringle IA, Sumner-Jones SG, Baker A, Tennant P, Gordon C, Vrettou C, Blundell R, Hyndman L, Stevenson B, Wilson A, Doherty A, Shaw DJ, Coles RL, Painter H, Cheng SH, Scheule RK, Davies JC, Innes**

JA, Hyde SC, Griesenbach U, Alton EWFW, Boyd AC, Porteous DJ, Gill DR, Collie DDS. 2011. Pre-clinical evaluation of three non-viral gene transfer agents for cystic fibrosis after aerosol delivery to the ovine lung. Gene Ther **18:**996-1005.

115. **Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, Doherty C, McLachlan G.** 2015. Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate sodium. PLoS ONE **10:**e0142097.

116. **Lu Q, Girardi C, Zhang M, Bouhemad B, Louchahi K, Petitjean O, Wallet F, Becquemin M-H, Le Naour G, Marquette C-H, Rouby J-J.** 2010. Nebulized and intravenous colistin in experimental pneumonia caused by *Pseudomonas aeruginosa*. Intensive Care Med **36:**1147-1155.

117. **Collie D, Govan J, Wright S, Thornton E, Tennant P, Smith S, Doherty C, McLachlan G.** 2013. A lung segmental model of chronic *Pseudomonas* infection in sheep. PLoS ONE **8:**e67677.

118. **Alley MR.** 1975. The bacterial flora of the respiratory tract of normal and pneumonic sheep. New Zeal Vet J **23:**113-118.

119. **Yimer N, Asseged B.** 2007. Aerobic bacterial flora of the respiratory tract of healthy sheep slaughtered in Dessie municipal abattoir, northeastern Ethiopia. Rev Med Vet Toulouse **158:**473-478.

120. **Barbour EK, Nabbut NH, Hamadeh SK, Al-Nakhli HM.** 1997. Bacterial identity and characteristics in healthy and unhealthy respiratory tracts of sheep and calves. Vet Res Commun **21:**421-430.

121. **Dunbar MR, Ward ACS, Power G.** 1990. Isolation of *Pasteurella haemolytica* from tonsillar biopsies of Rocky Mountain Bighorn Sheep. J Wildl Dis **26:**210-213.

122. **Al-Tarazi YHM, Dagnall GJR.** 1997. Nasal carriage of *Pasteurella haemolytica* serotypes by sheep and goats in Jordan. Trop Anim Health Prod **29:**177-179.

123. **Queen C, Ward ACS, Hunter DL.** 1994. Bacteria isolated from nasal and tonsillar samples of clinically healthy Rocky Mountain bighorn and domestic sheep. J Wildl Dis **30:**1-7.

124. **Shreeve BJ, Thompson DA, Biberstein EL.** 1972. Variation in carrier rates of *Pasteurella-haemolytica* in sheep .2. Diseased flocks. J Comp Pathol **82:**111-116.

125. **Cutlip RC, Lehmkuhl HD, Brogden KA.** 1993. Chronic effects of coinfection in lambs with parainfluenza-3 virus and *Pasteurella haemolytica*. Small Rumin Res **11:**171-178.

126. **Al-Darraji AM, Cutlip RC, Lehmkuhl HD, Graham DL.** 1982. Experimental infection of lambs with Bovine Respiratory Syncytial Virus and *Pasteurella haemolytica*: Pathologic studies. Am J Vet Res **43:**224-229.

127. **Davies DH, Herceg M, Thurley DC.** 1982. Experimental infection of lambs with an adenovirus followed by *Pasteurella haemolytica*. Vet Microbiol **7:**369-381.

128. **Trigo FJ, Breeze RG, Liggitt HD, Evermann JF, Trigo E.** 1984. Interaction of Bovine Respiratory Syncytial Virus and *Pasteurella haemolytica* in the ovine lung. Am J Vet Res **45:**1671-1678.

129. **Dassanayake RP, Shanthalingam S, Herndon CN, Subramaniam R, Lawrence PK, Bavananthasivam J, Cassirer EF, Haldorson GJ, Foreyt WJ, Rurangirwa FR, Knowles DP, Besser TE, Srikumaran S.** 2010. *Mycoplasma ovipneumoniae* can predispose bighorn sheep to fatal *Mannheimia haemolytica* pneumonia. Vet Microbiol **145:**354-359.

130. **Dassanayake RP, Call DR, Sawant AA, Casavant NC, Weiser GC, Knowles DP, Srikumaran S.** 2010. *Bibersteinia trehalosi* inhibits the growth of *Mannheimia haemolytica* by a proximity-dependent mechanism. Appl Environ Microbiol **76:**1008-1013.

131. **Bavananthasivam J, Dassanayake RP, Kugadas A, Shanthalingam S, Call DR, Knowles DP, Srikumaran S.** 2012. Proximity-dependent inhibition of growth of *Mannheimia haemolytica* by *Pasteurella multocida*. Appl Environ Microbiol **78:**6683-6688.

132. **Allen JW, Viel L, Bateman KG, Rosendal S.** 1992. Changes in the bacterial-flora of the upper and lower respiratory tracts and bronchoalveolar lavage differential cell counts in feedlot calves treated for respiratory-diseases. Can J Vet Res-Rev Can Rech Vet **56:**177-183.

133. **Chakravorty S, Helb D, Burday M, Connell N, Alland D.** 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods **69:**330-339.

134. **Větrovský T, Baldrian P.** 2013. The variability of the 16S rRNA gene in bacterial genomes and Its consequences for bacterial community analyses. PLoS ONE **8:**e57923.

135. **Kumar PS, Brooker MR, Dowd SE, Camerlengo T.** 2011. Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. PLoS ONE **6:**e20956.

136. **Guo F, Ju F, Cai L, Zhang T.** 2013. Taxonomic precision of different hypervariable regions of 16S rRNA gene and annotation methods for functional bacterial groups in biological wastewater treatment. PLoS ONE **8:**e76185.

137. **Liu Z, DeSantis TZ, Andersen GL, Knight R.** 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. Nucleic Acids Res **36:**e120.

138. **D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N.** 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics **17:**1-20.

139. **Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ.** 2016. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathogens **8:**24.

140. **Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW.** 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol **12:**87.

141. **Biesbroek G, Sanders EAM, Roeselers G, Wang XH, Caspers MPM, Trzcinski K, Bogaert D, Keijser BJF.** 2012. Deep sequencing analyses of low density microbial communities: Working at the boundary of accurate microbiota detection. PLoS ONE **7:**e32942.

142. **Antony KM, Ma J, Mitchell KB, Racusin DA, Versalovic J, Aagaard K.** 2015. The preterm placental microbiome varies in association with excess maternal gestational weight gain. Am J Obstet Gynecol **212:**e1-e16.

143. **Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S.** 2016. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. Sci Rep **6:**23129.

144. **Branton WG, Lu JQ, Surette MG, Holt RA, Lind J, Laman JD, Power C.** 2016. Brain microbiota disruption within inflammatory demyelinating lesions in multiple sclerosis. Sci Rep **6:**37344.

145. **Shin H, Price K, Albert L, Dodick J, Park L, Dominguez-Bello MG.** 2016. Changes in the eye microbiota associated with contact lens wearing. mBio **7:**e00198-00116.

146. **Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, Leite R, Elovitz MA, Parry S, Bushman FD.** 2016. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. Microbiome **4:**1-11.

147. **Marsh RL, Kaestli M, Chang AB, Binks MJ, Pope CE, Hoffman LR, Smith-Vaughan HC.** 2016. The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. Microbiome **4:**1-18.

148. **Tamariz J, Voynarovska K, Prinz M, Caragine T.** 2006. The application of ultraviolet irradiation to exogenous sources of DNA in plasticware and water for the amplification of low copy number DNA. J Forensic Sci **51:**790-794.

149. **Champlot S, Berthelot C, Pruvost M, Bennett EA, Grange T, Geigl E-M.** 2010. An efficient multistrategy DNA decontamination procedure of PCR reagents for hypersensitive PCR applications. PLoS ONE **5:**e13042.

150. **Czurda S, Smelik S, Preuner-Stix S, Nogueira F, Lion T.** 2016. Occurrence of fungal DNA contamination in PCR reagents: Approaches to control and decontamination. J Clin Microbiol **54:**148-152.

151. **Humphrey B, McLeod N, Turner C, Sutton JM, Dark PM, Warhurst G.** 2015. Removal of contaminant DNA by combined UV-EMA treatment allows low copy number detection of clinically relevant bacteria using pan-bacterial real-time PCR. PLoS ONE **10:**e0132954.

152. **Klaschik S, Lehmann L, Raadts A, Hoeft A, Stuber F.** 2002. Comparison of different decontamination methods for reagents to detect low concentrations of bacterial 16S DNA by real-time-PCR. Mol Biotechnol **22:**231-242.

153. **Borst A, Box ATA, Fluit AC.** 2004. False-positive results and contamination in nucleic acid amplification assays: Suggestions for a prevent and destroy strategy. Eur J Clin Microbiol Infect Dis **23:**289-299.

154. **Lazarevic V, Gaïa N, Girard M, Schrenzel J.** 2016. Decontamination of 16S rRNA gene amplicon sequence datasets based on bacterial load assessment by qPCR. BMC Microbiol **16:**1-8.

155. **Yu GQ, Fadrosh D, Goedert JJ, Ravel J, Goldstein AM.** 2015. Nested PCR biases in interpreting microbial community structure in 16S rRNA gene sequence datasets. PLoS ONE **10:**e0132253.

156. **Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, Knights D, Beckman KB.** 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat Biotechnol **34:**942-949.

157. **Kennedy K, Hall MW, Lynch MDJ, Moreno-Hagelsieb G, Neufeld JD.** 2014. Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles. Appl Environ Microbiol **80:**5717-5722.

158. **Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, Serrano MG, Reris RA, Sheth NU, Huang B, Girerd P, Strauss JF, Jefferson KK, Buck GA, Vaginal Microbiome C.** 2015. The truth about metagenomics: Quantifying and counteracting bias in 16S rRNA studies. BMC Microbiol **15:**66.

159. **Fouhy F, Clooney AG, Stanton C, Claesson MJ, Cotter PD.** 2016. 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. BMC Microbiol **16:**1-13.

160. **Sarkar G, Sommer SS.** 1993. Removal of DNA contamination in polymerase chain reaction reagents by ultraviolet irradiation, p 381-388. *In* Wu R (ed), Methods in Enzymology; Recombinant DNA, Part I, vol 218. Academic Press, Inc., London, UK.

161. **Cuthbertson L, Rogers GB, Walker AW, Oliver A, Hafiz T, Hoffman LR, Carroll MP, Parkhill J, Bruce KD, van der Gast CJ.** 2014. Time between collection and storage significantly influences bacterial sequence composition in sputum samples from cystic fibrosis respiratory infections. J Clin Microbiol **52:**3011-3016.

162. **Hang J, Desai V, Zavaljevski N, Yang Y, Lin XX, Satya RV, Martinez LJ, Blaylock JM, Jarman RG, Thomas SJ, Kuschner RA.** 2014. 16S rRNA gene pyrosequencing of reference and clinical samples and investigation of the temperature stability of microbiome profiles. Microbiome **2:**31.

163. **Zhao JC, Li J, Schloss PD, Kalikin LM, Raymond TA, Petrosino JF, Young VB, LiPuma JJ.** 2011. Effect of sample storage conditions on culture-independent bacterial community measures in cystic fibrosis sputum specimens. J Clin Microbiol **49:**3717-3718.

164. **Kwambana BA, Mohammed NI, Jeffries D, Barer M, Adegbola RA, Antonio M.** 2011. Differential effects of frozen storage on the molecular detection of bacterial taxa that inhabit the nasopharynx. BMC Clin Pathol **11:**1-8.

165. **Bahl MI, Bergstrom A, Licht TR.** 2012. Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. FEMS Microbiol Lett **329:**193-197.

166. **Falagas ME, Kasialkou SK.** 2006. Use of international units when dosing colistin will help decrease confusion related to various formulations of the drug around the world. Antimicrob Agents Chemother **50:**2274-2275.

167. **Li J, Nation RL, Turnidge JD.** 2006. Defining the dosage units for colistin methanesulfonate: Urgent need for international harmonization. Antimicrob Agents Chemother **50:**4231-4231.

168. **Michalopoulos A, Falagas M.** 2011. Colistin: Recent data on pharmacodynamics properties and clinical efficacy in critically ill patients. Ann Intensive Care **1:**30.

169. **Forest Laboratories UK Limited.** 2016. Colomycin Injection: Package leaflet: Information for the user.

170. **Dalfino L, Puntillo F, Mosca A, Monno R, Spada ML, Coppolecchia S, Miragliotta G, Bruno F, Brienza N.** 2012. High-dose, extended-interval colistin administration in critically ill patients: Is this the right dosing strategy? A preliminary study. Clin Infect Dis **54:**1720-1726.

171. **He H, Li JC, Nation RL, Jacob J, Chen G, Lee HJ, Tsuji BT, Thompson PE, Roberts K, Velkov T, Li J.** 2013. Pharmacokinetics of four different brands of colistimethate and formed colistin in rats. J Antimicrob Chemother **68:**2311-2317.

172. **Kiem S, Schentag JJ.** 2008. Interpretation of antibiotic concentration ratios measured in epithelial lining fluid. Antimicrob Agents Chemother **52:**24-36.

173. **Li J, Milne RW, Nation RL, Turnidge JD, Coulthard K, Johnson DW.** 2001. A simple method for the assay of colistin in human plasma, using pre-column derivatization with 9-fluorenylmethyl chloroformate in solid-phase extraction cartridges and reversed-phase high-performance liquid chromatography. J Chromatogr B Biomed Sci Appl **761:**167-175.

174. **Li J, Milne RW, Nation RL, Turnidge JD, Coulthard K, Valentine J.** 2002. Simple method for assaying colistin methanesulfonate in plasma and urine using high-performance liquid chromatography. Antimicrob Agents Chemother **46:**3304-3307.

175. **Ali SZ, Yang GCH.** 2000. Lung and Mediastinum Cytohistology. Cambridge University Press, Cambridge, UK.

176. **Willner D, Daly J, Whiley D, Grimwood K, Wainwright CE, Hugenholtz P.** 2012. Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. PLoS ONE **7:**e34605.

177. **Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM, UK IBD Genetics Consortium, Satsangi J, Flint HJ, Parkhill J, Lees CW, Hold GL.**

2014. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. PLoS ONE **9:**e88982.

178. **Yuan SQ, Cohen DB, Ravel J, Abdo Z, Forney LJ.** 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS ONE **7:**e33865.

179. **Abusleme L, Hong BY, Dupuy AK, Strausbaugh LD, Diaz PI.** 2014. Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing. J Oral Microbiology **6:**23990.

180. **Schloss PD.** 2010. The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. PLoS Comput Biol **6:**e1000844.

181. **Berry D, Ben Mahfoudh K, Wagner M, Loy A.** 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl Environ Microbiol **77:**7846-7849.

182. **Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ.** 2012. Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol **30:**562-562.

183. **Schirmer M, D'Amore R, Ijaz UZ, Hall N, Quince C.** 2016. Illumina error profiles: Resolving fine-scale variation in metagenomic sequencing data. BMC Bioinform **17:**1-15.

184. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnetjournal **17:**10-12.

185. **Jiang HS, Lei R, Ding SW, Zhu SF.** 2014. Skewer: A fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics **15:**182.

186. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:**7537-7541.

187. **Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD.** 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Appl Environ Microbiol **79:**5112-5120.

188. **Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN.** 2015. A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. J Proteomics Bioinform **8:**283-291.

189. **Nilakanta H, Drews KL, Firrell S, Foulkes MA, Jablonski KA.** 2014. A review of software for analyzing molecular sequences. BMC Res Notes **7:**830.

190. **Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO.** 2007. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res **35:**7188-7196.

191. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R.** 2011. UCHIME improves sensitivity and speed of chimera detection. BMC Bioinform **27:**2194-2200.

192. **DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL.** 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72:**5069-5072.

193. **Werner JJ, Koren O, Hugenholtz P, DeSantis TZ, Walters WA, Caporaso JG, Angenent LT, Knight R, Ley RE.** 2012. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. ISME J **6:**94-103.

194. **Balvočiūtė M, Huson DH.** 2017. SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? BMC Genomics **18:**114.

195. **Good IJ.** 1953. The population frequencies of species and the estimation of population parameters. Biometrika **40:**237-264.

196. **Magurran AE.** 1988. Ecological Diversity and its Measurement. Princeton University Press Princeton, US.

197. **Colwell RK, Coddington JA.** 1994. Estimating terrestrial biodiversity through extrapolation. Philos Trans R Soc Lond Ser B-Biol Sci **345:**101-118.

198. **Yue JC, Clayton MK.** 2005. A similarity measure based on species proportions. Commun Stat-Theory Methods **34:**2123-2131.

199. **Martin AP.** 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. Appl Environ Microbiol **68:**3673-3682.

200. **Schloss PD.** 2008. Evaluating different approaches that test whether microbial communities have the same structure. Isme J **2:**265-275.

201. **Holmes I, Harris K, Quince C.** 2012. Dirichlet multinomial mixtures: Generative models for microbial metagenomics. PLoS ONE **7:**e30126.

202. **White JR, Nagarajan N, Pop M.** 2009. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol **5:**e1000352.

203. **Dufrene M, Legendre P.** 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. Ecol Monogr **67:**345-366.

204. **R Core Team.** 2015. R: A language and environment for statistical computing, R Foundation for Statistical Computing, https://www.R-project.org/.

205. **Warnes G, Bolker B, Bonebakker L, Gentleman R, Liaw W, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B.** 2015. gplots: Various R programming tools for plotting data, v2.17.0. http://CRAN.R-project.org/package=gplots.

206. **Ploner A.** 2015. Heatplus: Heatmaps with row and/or column covariates and colored clusters, v2.16.0. https://github.com/alexploner/Heatplus.

207. **Neuwirth E.** 2014. RColorBrewer: ColorBrewer palettes, v1.1-2. http://CRAN.R-project.org/package=RColorBrewer.

208. **Oksanen J, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Simpson G, Solymos P, Stevens M, Wagner H.** 2016. Vegan: Community ecology package, v2.3-4/r2988. http://R-Forge.R-project.org/projects/vegan/}.

209. **Bray JR, Curtis JT.** 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol Monograph **27:**326-349.

210. **Martinsen EMH, Leiten EO, Bakke PS, Eagan TML, Gronseth R.** 2016. Participation in research bronchoscopy: A literature review. Eur Respir J **3:**29511.

211. **Yadava K, Pattaroni C, Sichelstiel AK, Trompette A, Gollwitzer ES, Salami O, von Garnier C, Nicod LP, Marsland BJ.** 2015. Microbiota promotes chronic pulmonary inflammation by enhancing IL-17A and autoantibodies. Am J Respir Crit Care Med **115:**63-68.

212. **Yu W, Yuan X, Xu X, Ding R, Pang L, Liu Y, Guo Y, Li H, Li M, Yuan J, Tang L, Wen S.** 2015. Reduced airway microbiota diversity is associated with elevated allergic respiratory inflammation. Ann Allergy Asthma Immunol **115:**63-68.

213. **Stokes K, Wood A, August A.** 2015. The microbiome regulates the eosinophil requirement for development of allergic lung inflammation. J Immunol **194:**53.18.

214. **Barfod KK, Roggenbuck M, Hansen LH, Schjørring S, Larsen ST, Sørensen SJ, Krogfelt KA.** 2013. The murine lung microbiome in relation to the intestinal and vaginal bacterial communities. BMC Microbiol **13:**1-12.

215. **Barfod KK, Vrankx K, Mirsepasi-Lauridsen HC, Hansen JS, Hougaard KS, Larsen ST, Ouwenhand AC, Krogfelt KA.** 2015. The murine lung microbiome changes during lung inflammation and intranasal vancomycin treatment. Open Microbiol J **9:**167-179.

216. **Singanayagam A, Glanville N, Pearson R, James P, Cuthbertson L, Cox M, Moffatt M, Cookson W, Bartlett N, Johnston S.** 2015. T1 Fluticasone propionate alters the resident airway microbiota and impairs anti-viral and anti-bacterial immune responses in the airways. Thorax **70:**A1.

217. **Bouljihad M, Leipold HW.** 1994. An ultrastructural study of pulmonary bronchiolar and alveolar epithelium in sheep. J Vet Med A **41:**573-586.

218. **Griebel PJ, Entrican G, Rocchi M, Beskorwayne T, Davis WC.** 2007. Cross-reactivity of mAbs to human CD antigens with sheep leukocytes. Vet Immunol Immunopathol **119:**115-122.

219. **Esty WW.** 1986. The efficiency of Good's nonparametric coverage estimator. Ann Stat **14:**1257-1260.

220. **Dufrêne M, Legendre P.** 1997. Species assemblages and indicator species:The need for a flexible assymetrical approach. Ecol Monogr **67:**345-366.

221. **Entrican G, Wattegedera SR, Griffiths DJ.** 2015. Exploiting ovine immunology to improve the relevance of biomedical models. Mol Immunol **66:**68-77.

222. **Enkhbaatar P, Nelson C, Salsbury JR, Carmical JR, Torres KEO, Herndon D, Prough DS, Luan L, Sherwood ER.** 2015. Comparison of gene expression by sheep and human blood stimulated with the TLR4 agonists lipopolysaccharide and monophosphoryl lipid A. PLoS One **10:**e0144345.

223. **Huws SA, Kim EJ, Lee MRF, Scott MB, Tweed JKS, Pinloche E, Wallace RJ, Scollan ND.** 2011. As yet uncultured bacteria phylogenetically classified as *Prevotella*, Lachnospiraceae *incertae sedis* and unclassified Bacteroidales, Clostridiales and Ruminococcaceae may play a predominant role in ruminal biohydrogenation. Environ Microbiol **13:**1500-1512.

224. **Zened A, Combes S, Cauquil L, Mariette J, Klopp C, Bouchez O, Troegeler-Meynadier A, Enjalbert F.** 2013. Microbial ecology of the rumen evaluated by 454 GS FLX pyrosequencing is affected by starch and oil supplementation of diets. FEMS Microbiol Ecol **83:**504-514.

225. **Omoniyi LA, Jewell KA, Isah OA, Neumann AP, Onwuka CFI, Onagbesan OM, Suen G.** 2014. An analysis of the ruminal bacterial microbiota in West African Dwarf sheep fed grass- and tree-based diets. J Appl Microbiol **116:**1094-1105.

226. **Christensen H, Kuhnert P, Nørskov-Lauritsen N, Planet PJ, Bisgaard M.** 2014. The Family Pasteurellaceae, p 535-564. *In* Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The Prokaryotes: Gammaproteobacteria. Springer Berlin Heidelberg, Berlin, Heidelberg.

227. **Ling Z, Liu X, Wang Y, Li L, Xiang C.** 2012. Pyrosequencing analysis of the salivary microbiota of healthy Chinese children and adults. Microb Ecol **65:**487-495.

228. **Cardenas PA, Cooper PJ, Cox MJ, Chico M, Arias C, Moffatt MF, Cookson WO.** 2012. Upper airways microbiota in antibiotic-naïve wheezing and healthy infants from the tropics of rural Ecuador. PLoS One **7:**e46803.

229. **May NDS.** 1964. The Anatomy of the Sheep. St. Lucia: University of Queensland Press, Queensland, Australia.

230. **Hungate RE.** 1966. Chapter V - The Rumen as a Continuous Fermentation System, p 206-244, The Rumen and its Microbes. Academic Press, NY, USA.

231. **Sarkar A, Singh H.** 2012. Oral Behaviour of Food Emulsions, p 111-137, Food Oral Processing: Fundamentals of Eating and Sensory Perception. Wiley-Blackwell, Online.

232. **Athanassa ZE, Markantonis SL, Fousteri MZF, Myrianthefs PM, Boutzouka EG, Tsakris A, Baltopoulos GJ.** 2012. Pharmacokinetics of inhaled colistimethate sodium (CMS) in mechanically ventilated critically ill patients. Intensive Care Med **38:**1779-1786.

233. **Hindler JA, Humphries RM.** 2013. Colistin MIC variability by method for contemporary clinical isolates of multidrug-resistant Gram-negative bacilli. J Clin Microbiol **51:**1678-1684.

234. **Quinn PJ, Markey BK, Leonard FC, Hartigan P, Fanning S, Fitzpatrick ES.** 2011. Chapter 14: *Staphylococcus* Species. Wiley-Blackwell, UK.

235. **Timsit E, Workentine M, Schryvers AB, Holman DB, van der Meer F, Alexander TW.** 2016. Evolution of the nasopharyngeal microbiota of beef cattle from weaning to 40 days after arrival at a feedlot. Vet Microbiol **187:**75-81.

236. **Hoyles L, Collins MD, Foster G, Falsen E, Schumann P.** 2004. *Jeotgalicoccus pinnipedialis* sp nov., from a southern elephant seal (*Mirounga leonina*). Int J Syst Evol Microbiol **54:**745-748.

237. **Misic AM, Davis MF, Tyldsley AS, Hodkinson BP, Tolomeo P, Hu BF, Nachamkin I, Lautenbach E, Morris DO, Grice EA.** 2015. The shared microbiota of humans and companion animals as evaluated from *Staphylococcus* carriage sites. Microbiome **3:**2.

238. **Schwaiger K, Hölzel C, Mayer M, Bauer J.** 2010. Notes on the almost unknown genus *Jeotgalicoccus*. Lett Appl Microbiol **50:**441-444.

239. **Effros RM, Dunning MB, Biller J, Shaker R.** 2004. The promise and perils of exhaled breath condensates. Am J Physiol-Lung Cell Mol Physiol **287:**L1073-L1080.

240. **Bondesson E, Jansson LT, Bengtsson T, Wollmer P.** 2009. Exhaled breath condensate-site and mechanisms of formation. J Breath Res **3:**016005.

241. **May AK, Brady JS, Romano-Keeler J, Drake WP, Norris PR, Jenkins JM, Isaacs RJ, Boczko EM.** 2015. A pilot study of the noninvasive assessment of the lung microbiota as a potential tool for the early diagnosis of ventilator-associated pneumonia. Chest **147:**1494-1502.

242. **Zakharkina T, Koczulla AR, Mardanova O, Hattesohl A, Bals R.** 2011. Detection of microorganisms in exhaled breath condensate during acute exacerbations of COPD. Respirology **16:**932-938.

243. **Jain R, Schriever CA, Danziger LH, Cho SH, Rubinstein I.** 2007. The IS6110 repetitive DNA element of *Mycobacterium tuberculosis* is not detected in exhaled breath condensate of patients with active pulmonary tuberculosis. Respiration **74:**329-333.

244. **St George K, Fuschino ME, Mokhiber K, Triner W, Spivack SD.** 2010. Exhaled breath condensate appears to be an unsuitable specimen type for the detection of influenza viruses with nucleic acid-based methods. J Virol Methods **163:**144-146.

245. **Vogelberg C, Hirsch T, Rosen-Wolff A, Kerkmann ML, Leupold W.** 2003. *Pseudomonas aeruginosa* and *Burkholderia cepacia* cannot be detected by PCR in the breath condensate of patients with cystic fibrosis. Pediatr Pulmonol **36:**348-352.

246. **Ritari J, Salojarvi J, Lahti L, de Vos WM.** 2015. Improved taxonomic assignment of human intestinal 16S rRNA sequences by a dedicated reference database. BMC Genomics **16:**1056.

Chapter 9: Appendices

Appendix 1: Lung microbiota changes associated with chronic *Pseudomonas aeruginosa* lung infection and the impact of intravenous colistimethate sodium

PLOS | ONE

RESEARCH ARTICLE

Lung Microbiota Changes Associated with Chronic *Pseudomonas aeruginosa* Lung Infection and the Impact of Intravenous Colistimethate Sodium

David Collie[1]*, Laura Glendinning[1], John Govan[2], Steven Wright[1], Elisabeth Thornton[1], Peter Tennant[1], Catherine Doherty[2], Gerry McLachlan[1]

1 The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, United Kingdom, **2** University of Edinburgh, Medical School, Edinburgh, Scotland, United Kingdom

* david.collie@ed.ac.uk

Citation: Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, et al. (2015) Lung Microbiota Changes Associated with Chronic *Pseudomonas aeruginosa* Lung Infection and the Impact of Intravenous Colistimethate Sodium. PLoS ONE 10 (11): e0142097. doi:10.1371/journal.pone.0142097

Editor: Mohamed N. Seleem, Purdue University, UNITED STATES

Received: June 10, 2015

Accepted: October 16, 2015

Published: November 6, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by the Biological and Biotechnological Science Research Council (BBSRC; www.bbsrc.ac.uk). Edinburgh Genomics is partly supported through core grants from NERC (R8/H10/56), MRC (MR/K001744/1) and BBSRC (BB/J004243/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Background

Exacerbations associated with chronic lung infection with *Pseudomonas aeruginosa* are a major contributor to morbidity, mortality and premature death in cystic fibrosis. Such exacerbations are treated with antibiotics, which generally lead to an improvement in lung function and reduced sputum *P. aeruginosa* density. This potentially suggests a role for the latter in the pathogenesis of exacerbations. However, other data suggesting that changes in *P. aeruginosa* sputum culture status may not reliably predict an improvement in clinical status, and data indicating no significant changes in either total bacterial counts or in *P. aeruginosa* numbers in sputum samples collected prior to pulmonary exacerbation sheds doubt on this assumption. We used our recently developed lung segmental model of chronic *Pseudomonas* infection in sheep to investigate the lung microbiota changes associated with chronic *P. aeruginosa* lung infection and the impact of systemic therapy with colistimethate sodium (CMS).

Methodology/Principal Findings

We collected protected specimen brush (PSB) samples from sheep (n = 8) both prior to and 14 days after establishment of chronic local lung infection with *P aeruginosa*. Samples were taken from both directly infected lung segments (direct) and segments spatially remote to such sites (remote). Four sheep were treated with daily intravenous injections of CMS between days 7 and 14, and four were treated with a placebo. Necropsy examination at d14 confirmed the presence of chronic local lung infection and lung pathology in every direct lung segment.

The predominant orders in lung microbiota communities before infection were Bacillales, Actinomycetales and Clostridiales. While lung microbiota samples were more likely to share

Competing Interests: The authors have declared that no competing interests exist.

similarities with other samples derived from the same lung, considerable within- and between-animal heterogeneity could be appreciated.

Pseudomonadales joined the aforementioned list of predominant orders in lung microbiota communities after infection. Whilst treatment with CMS appeared to have little impact on microbial community composition after infection, or the change undergone by communities in reaching that state, when Gram negative organisms (excluding Pseudomonadales) were considered together as a group there was a significant decrease in their relative proportion that was only observed in the sheep treated with CMS. With only one exception the reduction was seen in both direct and remote lung segments. This reduction, coupled with generally increasing or stable levels of Pseudomonadales, meant that the proportion of the latter relative to total Gram negative bacteria increased in all bar one direct and one remote lung segment.

Conclusions/Significance

The proportional increase in Pseudomonadales relative to other Gram negative bacteria in the lungs of sheep treated with systemic CMS highlights the potential for such therapies to inadvertently select or create a niche for bacteria seeding from a persistent source of chronic infection.

Introduction

Pseudomonas aeruginosa is considered to be the most important pathogen in cystic fibrosis (CF), with up to 60% of adult patients infected (UK CF Registry Annual Data Report 2014 [1]), and is also frequently isolated from patients with bronchiectasis [2]. In CF, early infections with *P. aeruginosa* can be transient, and can clear spontaneously, but colonization with *P. aeruginosa* usually occurs by the time patients reach their teenage years. In the later stages of infection, there is an adaptive shift from free-swimming planktonic *P. aeruginosa* to a sessile biofilm mode involving mucoid alginate-producing variants of the original colonising strain [3]. This important and characteristic shift is associated with more frequent and more severe pulmonary exacerbations (PEs) that result in progressive decrements in lung function [4].

P. aeruginosa also dominates chronic infections in a proportion of patients with bronchiectasis [5] and in chronic obstructive pulmonary disease (COPD) [6] where there is an increasing association with acute exacerbations.

The factors linking chronic *P. aeruginosa* lung infection to PEs are currently unknown. Certainly studies indicating that there is a reduction of sputum *P. aeruginosa* density during antibiotic treatment for PE in CF patients—a change that correlates with an improvement in lung function [7], tend to support a primary role for *P aeruginosa*. However, other data suggesting that changes in *P. aeruginosa* sputum culture status may not reliably predict an improvement in clinical status [8], and data indicating no significant changes in either total bacterial counts or in *P. aeruginosa* numbers in sputum samples collected prior to pulmonary exacerbation [9] sheds doubt on the specific role of *P. aeruginosa* in PE. Such uncertainty has been added to by recent 16S ribosomal DNA sequencing data. In a recent study of fifteen CF patients followed through 21 pulmonary exacerbations, sputum *P. aeruginosa* numbers did not increase immediately prior to a PE in CF adults [10]. These findings bear comparison with those of Carmody et al (2013) who found that during PE in CF patients bacterial community diversity and

bacterial density in sputum samples did not change between baseline and exacerbation [11], and Price et al (2013) who similarly found that total and relative abundance of genera at the population level were remarkably stable for individual patients regardless of clinical status [12]. These studies indicate that there are no generalizable ecological 'signatures' of PE in this type of clinical sample.

Daniels et al (2013)[13] investigated the relative impact of antibiotics, used predominantly to target *P. aeruginosa* during acute exacerbations, on other non-pseudomonal species. The relative abundance of viable *P. aeruginosa* and non-pseudomonal species was determined in sputa from adult CF subjects in the days preceding an exacerbation, and during and after antibiotic therapy, by T-RFLP profiling. Overall, an increase in the relative abundance of *P. aeruginosa* was observed, with a decrease in the total number of species detected. They raised the possibility that, aside from the direct effect of systemic antimicrobials on *P. aeruginosa*, there is coincident impact of antimicrobials on the remaining community members such that unspecified changes to *P. aeruginosa* gene expression may occur as a result of changes in interspecies communication. The potential exists for such changes in gene expression to impact on virulence and/or persistence [14].

Much of our current perceptions relating to the pathogenesis of PEs are driven by such studies relying on sputum to monitor inflammatory cells, bacterial densities, volatiles, mucin and protein content of the airways [15]. However, sputum characteristics can be highly variable between subjects and even within apparently stable subjects over time–reflecting, at least in part, heterogeneity of pathology across different lung regions and/or the relative contribution of different regions to the final sputum volume. Therefore sputum at best represents an averaging process, and at worst provides a highly skewed view of lung physiology and pathophysiology. These limitations may critically undermine our ability to u0nderstand the way in which heterogeneous disease processes, and associated microbiota, trigger fulminant whole-organ PEs.

Large animal models provide the means to dissect the pathophysiology of lung disease at a local level. Motivated by the current dearth of information surrounding the pathogenesis of PEs and speculation over the role of respiratory microbiota in this regard we recently developed a novel ovine model of chronic local lung infection with *P. aeruginosa* [16]. Our objectives in this research were to develop an understanding of the local pulmonary and microbiota response to chronic local lung infection with *P. aeruginosa* and to characterise the way in which systemic antibiotic therapy impacts on this response. We selected colistimethate sodium (CMS) as our antibiotic of choice. CMS undergoes hydrolysis in aqueous solutions to form a complex mixture of derivatives, including colistin [17]. Colistin is a polymyxin antibiotic with activity against Gram-negative organisms including *P. aeruginosa*. Whilst its use declined in the 1970s over concerns about toxicity it has recently experienced a resurgence as a consequence of the rise in resistance of *P. aeruginosa* and *Acinetobacter spp* to extended-spectrum cephalosporins, aminoglycosides, and fluoroquinolones. It is therefore a relevant choice in this context.

Methods

Ethics Statement

All experimental protocols were reviewed and approved by the local Roslin Institute ethical review process (The Roslin Institute Animal Welfare and Ethics Committee) and were subject to the provisions of the Animals (Scientific Procedures) Act 1986. During the course of the experimental protocols the animals were assessed on a daily basis for any clinical signs of adverse effect including dullness, depression, inappetence, coughing and/or dyspnoea.

Animals

Eight Suffolk cross sheep (4F & 4MN; Bodyweight 40.5 [38–48] Kg (Median[Range])) were used in this study. These sheep were commercially sourced and housed in groups on straw bedding under standard management conditions appropriate to a research setting. Sheep were randomly assigned to one of two treatment groups (Placebo (n = 4) and colistimethate sodium (CMS)(n = 4)).

Experimental Design

A baseline examination was conducted in which each sheep was subject to clinical examination and a blood sample taken from the jugular vein for routine haematological analysis. Thereafter each sheep was anaesthetised to facilitate bronchoscopic examination and sample collection according to standard protocol [16]. During this examination lung health was confirmed in the form of direct visualisation of the airway tree and later cytological analysis of bronchoalveolar lavage fluid (BALF). Protected specimen brush (PSB) samples were collected from the segmental bronchus serving the right apical (RA) lobe, the first ventral diaphragmatic (RVD1) segment of the right caudal diaphragmatic lobe, and the left cardiac (LC) segment of the apicocardiac lobe (Fig 1). BALF was subsequently collected from RA. After a recovery period of not less than two weeks (23 [15–34] days), the sheep were re-anaesthetised and Pseudomonas in agar beads ($2.5x10^9$ cfu in 2.5ml) instilled into the right cardiac (RC) lobe, the second ventral diaphragmatic (RVD2) segment of the right caudal diaphragmatic lobe, the left cardiac (LC) segment of the apicocardiac lobe, and the second ventral diaphragmatic (LVD2) segment of the left caudal diaphragmatic lobe. The method of instillation followed that previously described [16]. Three days later the sheep were blood sampled, anaesthetised again, and these instillations repeated. Four days after the second instillation the sheep assigned to the CMS group commenced daily treatment with intravenously administered systemic antibiotic (an intravenous dose of 50,000 international units (IU) kg-1 of colistimethate sodium (Colomycin®, Forest Laboratories UK Ltd, Dartford, Kent) every 24 h), with the placebo group commencing daily injections of saline. Eleven days after the second instillation–after one week of daily injections—blood samples were acquired before the sheep were anaesthetised and PSB specimens obtained from the previously sampled lobes (RA, RVD1 and LC). The sheep were then euthanized by intravenous injection of baribiturate, and the heart and lungs carefully removed from the carcase following standard necropsy protocols before the heart was dissected away and the lungs presented for further sampling and analysis. BALF was derived from each segment under study (RA, RC, RVD1, RVD2, LC and LVD2) prior to further dissection, sampling and recording using previously described methodology [16]. This experimental protocol and sampling scheme therefore allowed us to evaluate the effects of local lung infection with P aer., both within the direct segments (RC, LC, RVD2 and LVD2), within non-infected segments remote to the sites of direct infection (RA and RVD1), and systemically.

Anaesthesia

Food was withheld for 12 hours prior to anaesthesia. General anaesthesia was induced with intravenously injected propofol (6–8mg/kg)(Fresenius propofol, 1%, Fresenius Kabi Ltd) and anaesthesia maintained using positive pressure ventilation (Model 708; Harvard Apparatus, Millis, MA) with a 2:1 mixture of oxygen and nitrous oxide, and 1–3% isoflurane. Tidal volume was adjusted to 10ml/kg bodyweight and respiratory rate set to maintain end-tidal CO2 in the range 4.5–5.5%.

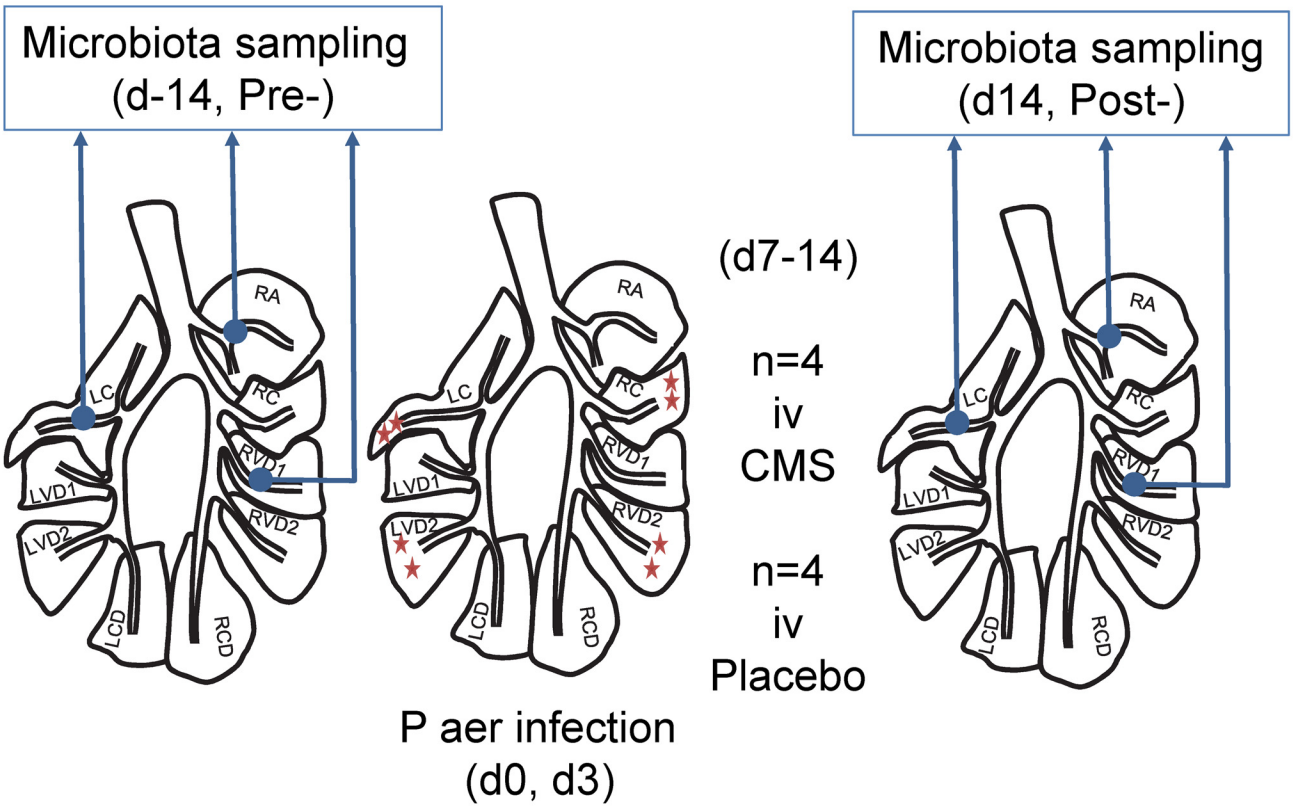


Fig 1. Microbiota sampling protocol. During a baseline examination lung health was confirmed in the form of direct visualisation of the airway tree and later cytological analysis of bronchoalveolar lavage fluid (BALF). Protected specimen brush (PSB) samples were collected from the segmental bronchus serving the right apical (RA) lobe, the first ventral diaphragmatic (RVD1) segment of the right caudal diaphragmatic lobe, and the left cardiac (LC) segment of the apicocardiac lobe (Pre-samples). BALF was subsequently collected from RA. At least two weeks later *P. aeruginosa* agar beads were instilled into the right cardiac (RC) lobe, the second ventral diaphragmatic (RVD2) segment of the right caudal diaphragmatic lobe, the left cardiac (LC) segment of the apicocardiac lobe, and the second ventral diaphragmatic (LVD2) segment of the left caudal diaphragmatic lobe. Three days later these instillations were repeated. Four days after the second instillation sheep were randomly assigned to daily intravenous injections of either CMS or placebo. Eleven days after the second instillation–after one week of daily injections—PSB specimens were obtained from the previously sampled lobes (RA, RVD1 and LC) (Post-samples). After the sheep was killed and the lungs removed for further analysis, BALF was derived from each segment under study (RA, RC, RVD1, RVD2, LC and LVD2) prior to further dissection, sampling and recording.

doi:10.1371/journal.pone.0142097.g001

P. aeruginosa Culture and Bead Preparation

P. aeruginosa embedded beads were prepared according to previously published methodology [16]. Briefly, broth cultures of *P. aeruginosa* mucoid strain PA0579 were prepared and cell suspensions mixed with molten agar before being injected into rapidly stirred heavy mineral oil. Beads were recovered thereafter by centrifugation.

Bronchial Brush Biopsy

A bronchoscope (FG-16X; Pentax, Englewood, CO, USA) was advanced down the trachea and mainstem bronchi and then into the relevant segmental bronchi until the predefined area selected for brushing was identified. The protected specimen brush (ConMed Endoscopic Technologies, Disposable Microbiology Brush 130) was then advanced through the biopsy channel of the bronchoscope. The plug was expelled and the sheath was retracted before the brush was applied to the epithelial mucosal surface. By advancing and retracting the brush in

contact with the mucosa, a sample of epithelial lining cells and fluid was obtained. The bronchial brush biopsy (BBr) sample was taken to a sterile flow cabinet and the brush end cut off directly into a non-stick Rnase free 1.5ml microfuge tube containing 1ml sterile phosphate buffered saline (Sigma D8537) and stored on ice. BBr samples were vortexed then the brush was removed under sterile conditions. Samples were centrifuged at 13000g for 15 minutes at 4°C and the pellet stored at -80°C.

Bronchoalveolar Lavage

The bronchoscope was wedged in selected segmental bronchi. Two 20ml aliquots of PBS (Sigma D8537) were used to collect BALF from selected lung segments. BALF samples were placed into sterile Falcon tubes and immediately placed on ice until subsequent analysis. BALF was centrifuged at 400g for seven minutes to separate out the cellular fraction and the resultant pellet was re-suspended in 2ml sterile PBS. The total cell number was counted before subsequent preparation of cytospins for differential cytology. Cells were counted using a Neubauer haemocytometer and values expressed per millilitre BALF. Cyto-centrifuge slides were prepared and stained using Leishman stain for differential counts on 500 cells. Cells were classified as neutrophils, macrophages, eosinophils, lymphocytes or mast cells according to standard morphological criteria. The remaining BALF was centrifuged at 13000g for 5 minutes at 4°C and the pellet and supernatant stored at -80°C.

Necropsy

Following euthanasia by intravenous injection of baribiturate, the heart and lungs were carefully removed from the carcase following standard necropsy protocols before the heart was dissected away and the lungs presented for further sampling and analysis. The trachea and right and left major bronchi were carefully opened along their dorsal aspect to expose the primary bronchial entrance to each lung segment of interest. Sterile polyethylene tubes were inserted in turn, into each lung segment bronchus until a wedge point was achieved. Thereafter 40ml sterile PBS (Sigma D8537) was instilled recovered and handled following the same principles employed during bronchoalveolar lavage under anaesthesia. Control lung segments were always sampled prior to handling segments previously exposed to Pseudomonas and all efforts were directed towards minimising the potential for cross-contamination between lung segments. Lung segments were carefully isolated by gross dissection from surrounding lung tissue before being separately examined and further dissected by parallel transverse sectioning along the plane of the subsegmental bronchus into ~1cm thick lung slices. Photographic images of the lung slices were collected. Samples were collected for assessing the degree of Pseudomonas infection.

Pathology Grading

Photographic images of lung slices derived from each lung segment were assessed and scored for the presence (1) or absence (0) of the following gross pathological features–pleural oedema, pleural fibrosis, and the presence of fibrotic/granulomatous tissue or abscessation in the lung parenchyma. The cumulative score (range 0–4) for each slice image was then multiplied by the proportion, quantified using ImageJ ([18]), of the slice cut surface considered visibly abnormal to give a total pathology score for each segment. Scores were then normalised to a scale of zero (no pathology) to 100 (the most severe pathology) by dividing by the maximum score observed amongst all lung slices and multiplying by 100.

Pseudomonas Infection Level

Tissue samples were stored on ice and then finely chopped under sterile conditions. 300mg of tissue was weighed and placed into Lysing Matrix D tubes (MP Biomedicals 6913–500) containing 600μl sterile PBS. Samples were homogenised using a Fastprep FP120 (Thermo Electron) with 3 bursts of 20 seconds at setting 6.0 and 5 minute incubation on ice between each homogenisation step. Equal volumes of BALF samples were centrifuged for 10 minutes at 2700g and each pellet re-suspended in 1ml sterile ice-cold PBS. BALF pellets were homogenised as for the tissue samples. Bacterial load was assessed from all samples by preparation of 10 fold dilutions in ice-cold PBS and 100μl of chosen dilutions spread on Pseudomonas Isolation Agar (PIA)(Sigma) plates and incubated overnight at 37°C. Bacterial counts were calculated by manual counting of colony forming units, multiplication by the dilution factor, and a further 10 fold, to allow for 100μl inoculum to give a final count in cfu/ml.

Sensitivity

We assessed the sensitivity to CMS of the *P. aeruginosa* mucoid strain PA0579 used to infect the sheep, and isolates cultured from the infected lung tissue, using Etest strips (Etest, bioMérieux, France).

DNA Isolation

DNA was extracted by modification of a previously published method [19]. Briefly, BBr pellets were suspended in 60μls of solution C1, provided with the PowerSoil DNA Kit (PowerSoil DNA Isolation Kit, MO-BIO). This suspension was transferred into PowerSoil Bead Tubes along with 750μl of PowerSoil Bead Solution. Bead Tubes were heated at 65°C for 10 minutes to aid cell lysis then placed in a FastPrep FP120 Cell Disrupter for 45 seconds at 5.0m/sec. All further steps were carried out following the manufacturer's instructions except that the final elution volume was changed to 50μl rather than 100μl. Extracted DNA was stored at -80°C until used in a nested PCR with primers to amplify the hypervariable region V1-V4 of the 16S rRNA gene followed by bar coded primers to amplify the hypervariable V2-V3. The PCR products were Agencourt AMPure XP (Beckman Coulter) cleaned after each PCR run to remove smaller non-specific products and unused primers then the purified amplicon products from each sample were pyrosequenced.

Triplicate 20μl reactions were performed per sample using the LightCycler® 480 SYBR Green I Master mix (Roche), 1μl of extracted DNA solution and the 16S rRNA Q-PCR primers UniF340 (5'–ACTCCTACGGGAGGCAGCAGT–3') and UniR514 (5'–ATTACCGCGGCTG CTGGC–3') at a final concentration of 0.4μM. The following steps were performed: a pre-incubation step of 50°C (ramp rate: 4.80°C/s for 2 minutes) then 95°C (ramp rate: 4.80°C/ for 10 seconds) and an amplification step consisting of 45 cycles of 95°C (ramp rate: 4.80°C/s for 30 seconds) then 63°C (ramp rate: 2.50°C/s for 30 seconds). This was followed by a melting cycle consisting of 95°C (ramp rate: 4.80°C/s for 5 seconds) then 65°C (ramp rate: 4.80°C/s for 1 minute) followed by 97°C (ramp rate: 0.11°C/s, acquisition mode, continuous). A standard curve beginning at 1.7×10^9 copies in nuclease free water and continuing in 1:10 dilutions to 1.7×100 was generated from a PCR product obtained from the genomic DNA of *P. aeruginosa* mucoid strain PA0579 [20] using the 16S rRNA primers 28F and 805R (PCR method described below), in order to calculate the total 16S rRNA gene copy numbers.

Barcoding of 16S Amplicons and Pyrosequencing

All PCR steps used 25μls of PCR master mix (Q5® High-Fidelity 2X Master Mix, New England Biolabs) and 2.5μls of both forward and reverse primers. A nested protocol of two

rounds of PCR reactions was performed. A negative control was included in each PCR run consisting of nuclease free water.

The V1-V4 variable region of the bacterial 16S rDNA was amplified using the primers 28F (5'–GAGTTTGATCNTGGCTCAG–3') and 805R (5'–GACTACCAGGGTATCTAATC–3'). The conditions for the first round of PCR were: 94°C for 2 minutes followed by 30 cycles of 94°C for 1 minute, 55°C for 45 seconds and 72°C for 1.5 minutes followed by 72°C for 20 minutes.

The V2-V3 region was amplified using Truseq barcoded primers 104F (5'—GGACGGGT GAGTAACACGTG–3') and 519R (5'–GTNTTACNGCGGCKGCTG–3')[21]. The conditions for the second round of PCR were: 98°C for 30 seconds followed by 20 cycles of 98°C for 10 seconds, 67°C for 30 seconds and 72°C for 10 seconds followed by 72°C for 2 minutes. Amplicons from both rounds of PCR were purified using the AMPure XP system (AMPure XP PCR Purification, Agencourt).

Amplicons were sequenced using an Illumina MiSeq producing paired 250-nucleotide reads [22]. Two negative PCR controls were included in the sequencing run.

Data Analysis

Primers were removed using Cutadapt [23]. The MOTHUR program [24] was used for quality control and taxonomic assignment of reads, following a protocol developed for MiSeq by the MOTHUR creators [22]. Sequences were phylotyped using the SILVA reference alignment and any sequences which did not correctly align were removed. Chimeras were identified and removed using UCHIME [25] within MOTHUR. Sequences were taxonomically classified, using MOTHUR's Bayesian classifier, against the Greengenes database [26]. Quality control consisted of the removal of sequences if they were not assigned to bacteria; were identified as chimeric; contained ambiguous bases or homopolymers < 9 bases; did not align to the correct region of the 16S gene or were less than 359 bases long. All samples were found to have Good's coverage values greater than 0.99, indicating sequencing to sufficient depth for the purposes of this study.

For quality control purposes, water samples were sequenced and analyzed through the bioinformatics pipeline. These samples had a much less diverse microbial community composition with over 66% abundance accounted for by only three OTUs. Whilst these OTUs could also be found in lung microbiota samples their proportional abundance was typically much less with median [range] proportional abundances of 0.4 [0–6.5], 0.0 [0–14.5], and 0.0 [0–3.4] respectively. No specific adjustment in analysis was made for OTUs present in water control samples. Sequence data was submitted to the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra; accession number SRP064022).

P. aeruginosa Specific qPCR

Targeting the oprl gene was performed following a modification of a previously published method [27]. Briefly, 1μl DNA samples were run in triplicate with 12.5μl Qiagen Quantitect probe mastermix, 0.3μM of each OPRL primer and 0.2μM hydrolysis probe with the reaction made up to a final volume of 25μl with water. Gene copy number was calculated from a standard curve of genomic DNA of *P. aeruginosa* mucoid strain PA0579.

Statistical Approach

Where data was normally distributed parametric data analysis procedures were applied (t-test, ANOVA), otherwise non parametric alternatives (Mann Whitney Test, Kruskal Wallis Test, Wilcoxon Signed-Ranks test and Spearman rank correlation) were used. Heatmaps were

generated using the "heatmap.2" R package, version 2.10.1 (available at http://CRAN.R-project.org/package=gplots). Non-metric multidimensional scaling (NMDS) analysis applied to distance matrices of Bray-Curtis dissimilarities, and ordination plots were generated using the "vegan" R package. Permutational ANOVA (PERMANOVA) (using the R-vegan function adonis) on Bray-Curtis distance matrices facilitated analysis and partitioning of sums of squares.

Results

Clinical Response

There were no signs of adverse effect nor significant variation in bodyweight in any sheep during the experimental period. All rectal temperatures lay within the range 39.0–40.2 (normal range 37.9–40.3, n = 664; D Collie, unpublished observations) and there was no significant difference between the groups.

Routine Haematology

Instillation of P aer. was associated with a significant reduction in the total white blood cell count in peripheral blood measured three days after the first instillation, a reduction that reverted to baseline levels by fourteen days after the first instillation (ANOVA; p = 0.000, relative to baseline and to day 14). There was no difference between the groups in this respect. There was no significant change in any other measured haematological parameters.

Bronchoalveolar Lavage Cytology

Infection with P aer. was associated with an increase in bronchoalveolar cellularity. The predominant cell types involved comprised alveolar macrophages, neutrophils and lymphocytes. Bubble plots illustrating the relationship between the log-transformed absolute numbers of these cell types (in baseline and post-infection samples from directly infected and remote lung segments) and the level of infectivity in the same samples are depicted in S1 Fig. Cytospin images representative of baseline and post-infection cytology are also depicted in S2 Fig.

Bacterial Load in BAL and Tissue Specimens

P aer. infection was reliably detected in all direct lung segments by culture of BALF and/or lung tissue. Whilst BALF samples derived from direct lung segments had a significantly greater burden of P aer. than remote lung segments (Wilcoxon Signed Rank Test of median = 0 versus median > 0; P = 0.030))(Fig 2A) there was no significant difference between the placebo and CMS groups in this regard. Two BALF samples derived from remote lung segments (2S065_RA and 2S037_RVD1 (CMS)) demonstrated evidence of infection. We also examined tissue from lung segments for evidence of infection and again found that samples derived from direct lung segments had a significantly greater burden of P aer. than remote lung segments (Wilcoxon Signed Rank Test of median = 0 versus median > 0; P = 0.007)(Fig 2B) with no difference between the groups. Three tissue samples derived from remote lung segments (2S037_RVD1 (CMS) and 2D616_RA & RVD1 (Placebo)) demonstrated evidence of infection. The *P. aeruginosa* mucoid strain PA0579 used to infect the lung was sensitive to CMS, and we found no difference between the infecting strain, and isolates cultured from lung tissue obtained at necropsy, in this respect (data not shown). In addition colony morphotypes were, with one exception, uniform in appearance and mucoid in character. The exception concerned colonies grown from tissue samples derived from a directly infected lung segment (LVD2) of a

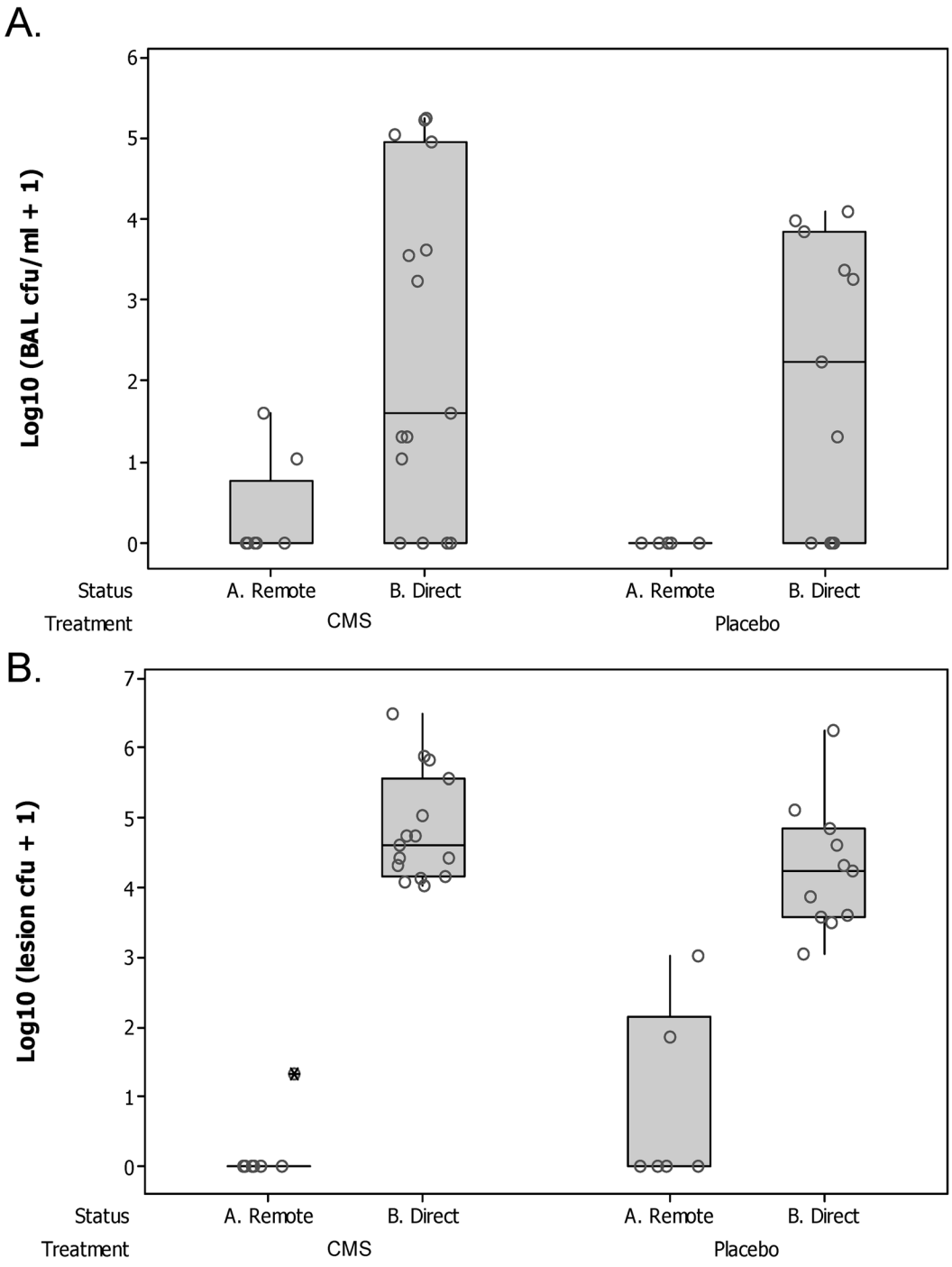


Fig 2. Bacterial burden in the lung. Boxplots showing the bacterial burden in A. bronchoalveolar lavage fluid (Log_{10}(BAL cfu/ml + 1) and B. lung tissue (Log_{10}(lesion cfu+1)) samples derived from lung segments directly infected 14 days previously with $2.5x10^9$ cfu *P. aeruginosa* in agar beads (B. Direct), and lung segments spatially remote to such segments (A. Remote). Sheep were treated with daily intravenous CMS (n = 4), or placebo (n = 4), between days 7–14. Bronchoalveolar lavage fluid samples derived from direct lung segments had a significantly greater burden of *P. aeruginosa* than remote lung segments (Wilcoxon Signed Rank Test of median = 0 versus median > 0; P = 0.030)); there was no significant difference between the placebo and CMS groups in this regard. Lung tissue samples derived from direct lung segments had a significantly greater burden of *P. aeruginosa* than remote lung segments (Wilcoxon Signed Rank Test of median = 0 versus median > 0; P = 0.007); there was no significant difference between the placebo and CMS groups in this regard.

doi:10.1371/journal.pone.0142097.g002

sheep treated with CMS (2D615). In this instance there was evidence of non-mucoid colonies intermixed with mucoid colonies.

Pathology

Semi-quantitative scoring of the gross pathological features indicated that there was a significant increase in gross pathology associated with direct infection with P aer. (Wilcoxon Signed Rank Test of median = 0 versus median > 0; P = 0.007) and that there was no difference between placebo- and CMS-treated sheep in this respect. Notably, with only one exception all the remote segments scored zero (the exception scoring 1), whereas the median [range] for the placebo sheep was 60.5 [12.3–95.0] and for the CMS sheep 30.5 [19.3–100.0]. There was a significant correlation between the pathology score and the burden of infection in BALF and tissue (Spearman Rho = 0.625 & 0.731, and P = 0.017 & 0.001 respectively).

Microbiota

The heatmap in Fig 3A illustrates that in the baseline (Pre) samples the most predominant consistently represented taxa were Bacillales (25±14% [1–49])(Mean±SD [Range]), Actinomycetales (18±9% [1–37]) and Clostridiales (14±10% [2–45]). Other taxa such as Enterobacteriales, Bacteroidales, Caulobacterales, Pasteurellales and Pseudomonadales also featured prominently but inconsistently.

For three sheep (2D616, 2S035 and 2D615) the predominant pattern was consistently represented in the samples derived from three different lung segments (Fig 4A), demonstrating within-lung homogeneity of microbiota. The remaining samples were found in sheep in which microbial communities were more diverse, demonstrating within-lung heterogeneity of microbiota. Visualisation of non-metric multi-dimensional scaling (NMDS) ordination analysis applied to the distance matrix of Bray-Curtis dissimilarities confirmed the above visual perceptions (Fig 5).

We performed hierarchical cluster analysis to ascertain whether samples would cluster by sheep and/or lung segment (Fig 3A) and employed permutational ANOVA (PERMANOVA) using the Bray-Curtis distance matrix to assess the extent to which variability could be assigned to the sheep, or lung segment, from which the communities were drawn. The results of the latter analysis indicated that a significant proportion of the variance (49%) could be explained by a sheep effect (P = 0.003), whereas the lung segment had no significant effect (P = 0.155).

The inverse Simpson index was calculated for each of the samples (Fig 3A). This index takes into account both species richness, and evenness of abundance among the species present. In essence it measures the probability that two entities taken at random from the dataset of interest represent the same type. Values ranged from 5.0 to 29.5, with an average of 14.1 and SD 6.25. There was no significant relationship between particular lung segments and the diversity of microbiota contained therein as measured by the Simpson Diversity Index (Kruskal Wallis Test, P = 0.735).

All sheep (n = 8) were then subjected to chronic local lung infection with *P. aeruginosa*. Four sheep were systemically treated with intravenous CMS and four treated with a placebo injection. The heatmap in Fig 3B illustrates that the four most predominant orders in the Post-samples consisted of Bacillales, Actinomycetales, Clostridiales and Pseudomonadales.

Fig 4B depicts the proportional representation of different phylotypes in post-infection (Post) samples derived from direct (LC) and remote (RA & RVD1) lung segments of each sheep and the relationship of these samples to their baseline (Pre) counterparts.

Hierarchical cluster analysis of the Post- samples failed to indicate clustering according to sheep, lung segment, treatment (placebo or CMS), or any change in diversity as reflected in the

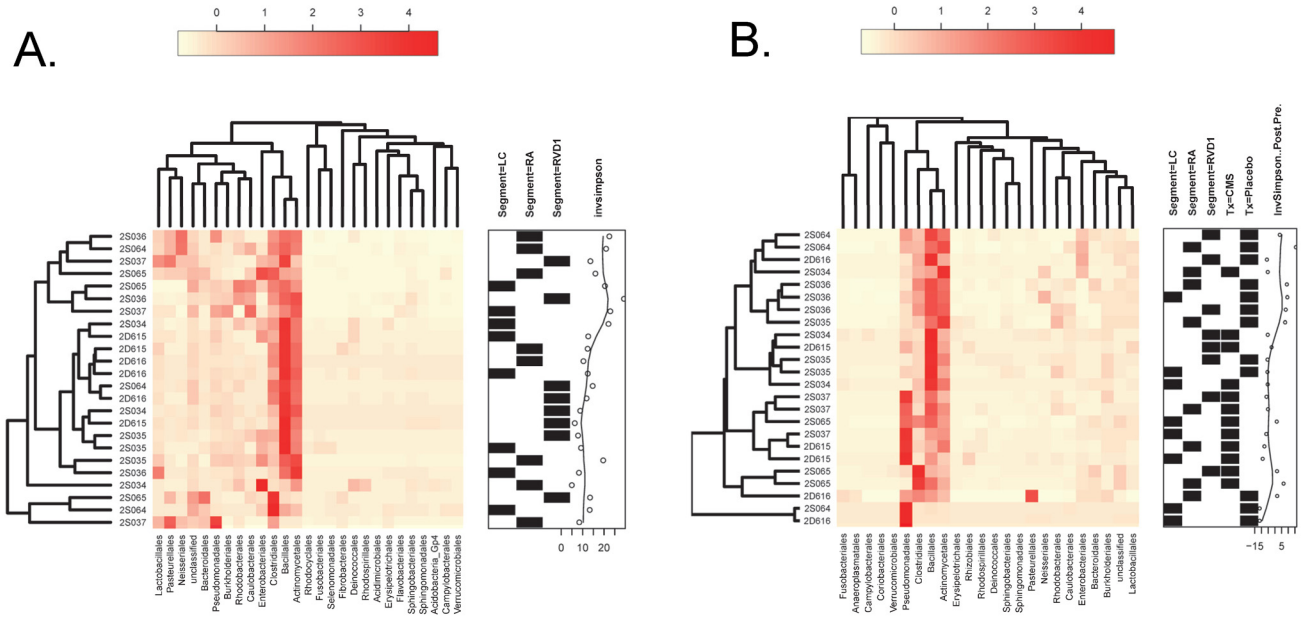


Fig 3. Abundance heatmaps for Pre- and Post-data. Heatmaps reflecting the proportional representation of microbiota (Order classification) in samples derived from different lung segments at baseline (A), and after lung infection and/or treatment with systemic CMS (B). The identities of individual sheep are indicated at the left side of each heatmap and the segment from which each sample was derived (LC, RA or RVD1) indicated in the annotation frame on the right side of the heatmap. Orders with a proportional representation of less than 1% are not shown. The results of hierarchical clustering applied to a distance matrix of Bray-Curtis measures between pairs of samples is shown for samples (left side) and bacteria (top). The remaining annotation in (A) reflects the inverse Simpson index that characterizes the species diversity in each sample community, with higher values reflecting an increase in diversity. In (B) the annotation frame also indicates whether the sample was derived from a sheep treated with CMS (Tx = CMS) or placebo (Tx = Placebo), and the change in the inverse Simpson index (Post-Pre) relative to the values at baseline, with positive values reflecting an increase, and negative values a decrease, in diversity.

doi:10.1371/journal.pone.0142097.g003

Post-Pre inverse Simpson Diversity index (Fig 3B). PERMANOVA analysis indicated that a significant proportion of the variance (42%) for Post- samples could be explained by a sheep effect (P = 0.01), and whether or not the segment had been directly infected (12%; P = 0.01). Whether or not the sheep had been treated with CMS or placebo had no significant bearing on the variance in the Post- samples.

We calculated the Log_2(Post-/Pre-) fold-change for the phylotypes in each set of paired samples (Fig 6) and examined the degree of relatedness between different samples by creating a distance matrix of Pearson correlation coefficients using the formula (1-Pearson Correlation Coefficient) as the index of dissimilarity. Hierarchical clustering failed to indicate any overarching influence of treatment, or lung segment, on the elicited patterns of change. PERMANOVA applied to a Euclidean distance matrix based on the sample-specific arrays of fold-changes failed to highlight any significant contributor to the variance observed.

We determined whether systemic therapy with CMS had any influence on the proportion of Gram negative bacteria (excluding Pseudomonadales) in lung microbiota. Whereas the presence of local lung infection with Ps aer did not significantly alter the proportion of Gram negative bacteria (excluding Pseudomonadales)(Fig 7A) in the lung segments of sheep treated with placebo (Paired t-test on Pre-Post differences (n = 4), P = 0.741), all of the direct segments and all except one of the remote segments (7/8) of sheep treated with CMS experienced a reduction in the proportion of Gram negative bacteria (excluding Pseudomonadales)(Fig 7B).

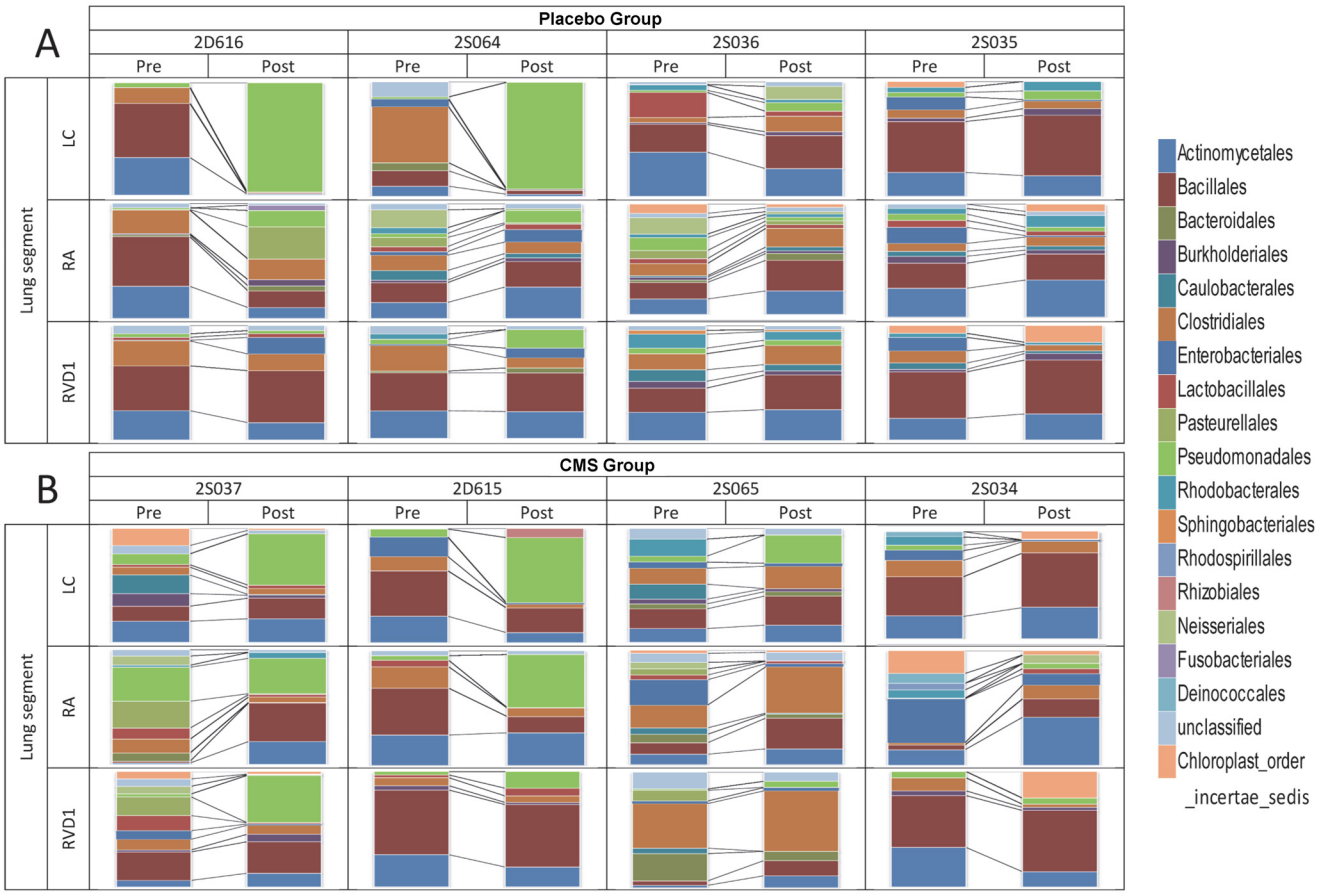


Fig 4. Proportional changes in microbiota from paired Pre- and Post- samples. Stacked column charts depicting the relative proportions of different bacterial phylotypes (classified at the level of Order, and coloured according to the legend) in PSB samples derived from three lung segments (Left cardiac (LC), Right apical (RA) and Right ventral diaphragmatic (RVD1)) prior to (Pre) and 14 days after (Post) the initiation of chronic lung infection with *P. aeruginosa* in segment LC. Four sheep (2D616, 2S064, 2S036 and 2S035) were treated with daily intravenous injections of saline (A. Placebo Group), and four sheep (2S037, 2D615, 2S065 and 2S034) were treated with daily intravenous injections of CMS (B. CMS Group) between days 7–14.

doi:10.1371/journal.pone.0142097.g004

The reduction in the proportion of G-ve bacteria (excluding Pseudomonadales) in lung segments of the CMS group was significant (Paired t-test on Pre-Post differences (n = 4), P = 0.040). When only the remote lung segments were included in this analysis, again the proportion of Gram negative bacteria (excluding Pseudomonadales) showed no significant change in response to infection in the placebo group whereas there was a significant reduction in this proportion in the sheep treated with CMS (Paired t-test on Pre-Post differences (n = 4), P = 0.979, and P = 0.044 respectively).

We determined, for the same samples, whether there was a relationship between qPCR for PA0579 and reads assigned to Pseudomonadales following 16SrRNA sequencing. There was a highly significant positive correlation (Spearmans rho = 0.711 (n = 19, $p<0.001$)).

Discussion

Instillation of *P. aeruginosa* in agar beads consistently induced a chronic local lung infection that resulted in gross pathology that remained confined to the areas where the instillate was

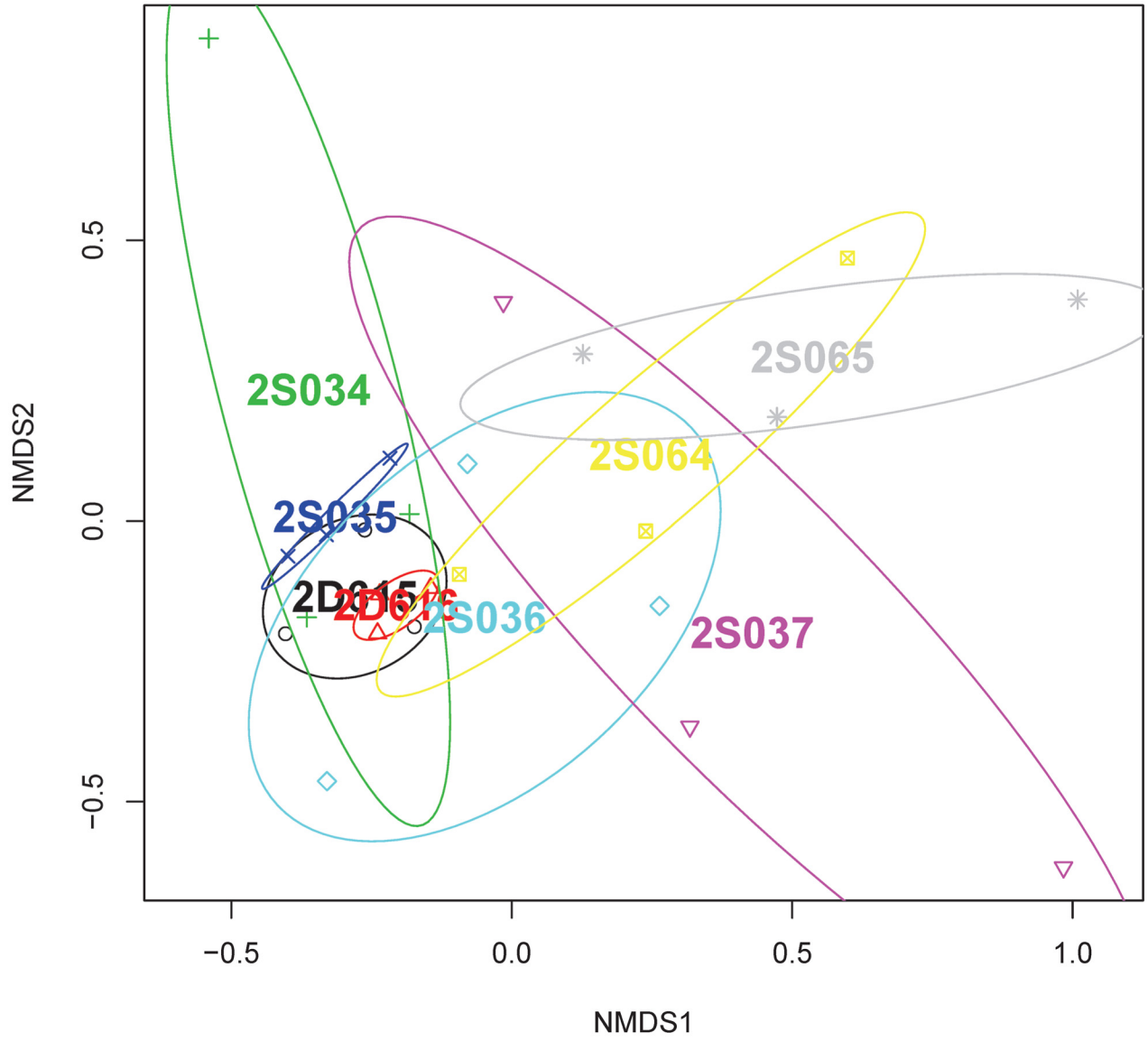


Fig 5. Non-metric multi-dimensional scaling (NMDS) ordination of bacterial communities in the lungs of healthy sheep. Individual plot data is grouped according to sheep identity and can be related through shared colour, a process facilitated by the coloured ellipses. Whilst some sheep have microbial communities that cluster tightly, and others are widely scattered, there is considerable overlap apparent between most of the sheep.

doi:10.1371/journal.pone.0142097.g005

delivered. The gross pathological features and changes in bronchoalveolar cytology were consistent with our previous experience with this model system [16].

We were able to assess the composition of lung microbiota across the lung by sampling from different airways of the same animals at the same point in time and to compare these samples amongst different animals.

We found that the microbial communities of the samples obtained at baseline were largely dominated by the orders Bacillales, Actinomycetales and Clostridiales. Whilst there was evidence for both within-lung homogeneity and heterogeneity amongst different animals, samples

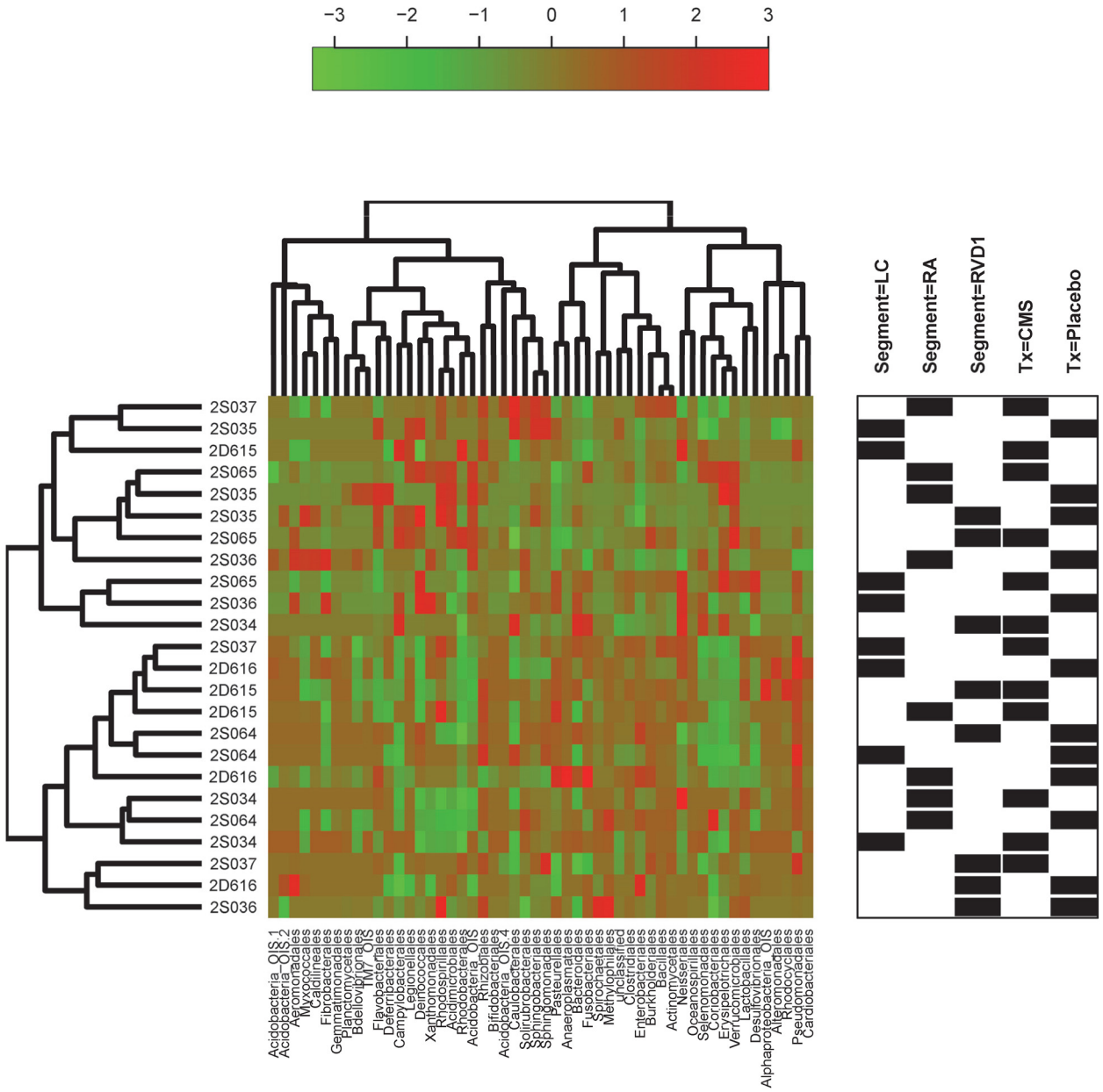


Fig 6. Fold change heatmap. Heat map representing the Log_2(Post-/Pre-) fold-change for the phylotypes in each set of paired samples. The degree of relatedness between different samples was assessed by creating a distance matrix of Pearson correlation coefficients using the formula (1-Pearson Correlation Coefficient) as the index of dissimilarity. Hierarchical clustering failed to indicate any overarching influence of treatment, or lung segment, on the elicited patterns of change.

doi:10.1371/journal.pone.0142097.g006

obtained from a given sheep were more likely to share similarity with other samples from the same sheep than from samples from different sheep.

Lung microbiota data derived from healthy human subjects is becoming increasingly available [19, 28, 29]. Charlson et al (2011) found that the predominant orders represented in BALF and PSB samples were Bacteroidales (dominated by Prevotella spp.), Clostridiales,

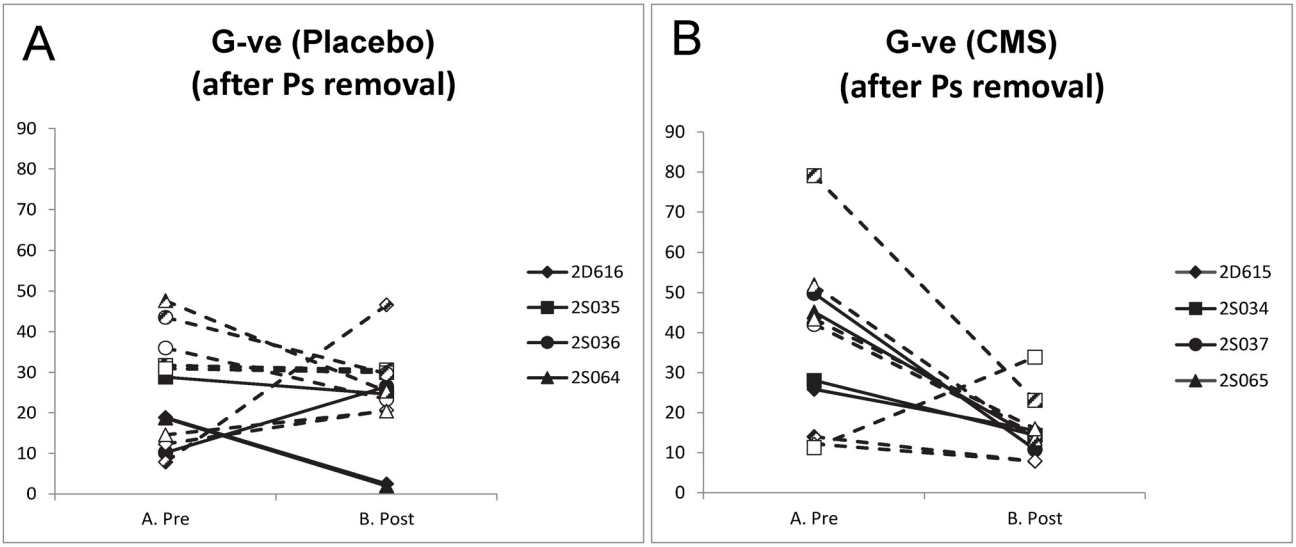


Fig 7. Change in Gram negative bacteria (excluding Pseudomonadales) in Pre- and Post-samples. Line charts indicating the percentage of Gram negative bacteria in PSB samples derived prior to (Pre) and 14 days after infection with P.aer (Post) from sheep treated with daily intravenous injections of A. saline (Placebo), or B. CMS (CMS) between days 7–14. Symbols reflect the identity of sheep according to the legend entries, with filled symbols reflecting samples derived from directly infected lung segments (LC), partially filled symbols from remote segment RA, and open symbols from remote segment RVD1.

doi:10.1371/journal.pone.0142097.g007

Lactobacillales (dominated by Streptococcal spp.) and Actinomycetales [19]. Hilty et al (2010) [30] similarly found that samples from healthy human subjects (adults and children) were dominated by Bacteroidales (Prevotella spp.), as did Dickson et al (2015) [29]. Indeed it is perhaps the prevalence of Bacteroidales, and consistent presence of Lactobacillales, that links these different studies of healthy human subjects. Otherwise, these separate studies are notable for their respective differences that are largely unexplained–such as the notable presence of Methylobacterium (Order: Rhizobiales) in the recent study of Dickson et al (2015) [29].

Bacteria of the order Baccillales, the most predominant members of the sheep lung microbiota, are therefore relatively infrequently found in human lungs whilst bacteria belonging to the other predominant orders, Actinomycetales and Clostridiales, can be found in human lungs but are inconsistently present.

The relatively minor contribution of Bacteroidales to ovine lung microbiota (~5%) is an interesting observation given both the predominance of Prevotella in the human lung and its high abundance in the rumen of sheep [31] where they help the breakdown of protein and carbohydrate foods. If measured lung microbiota in sheep are derived transiently from the oropharynx (itself heavily influenced by rumen contents through the process of rumination) then it would be reasonable to assume that Bacteriodales would feature more prominently.

Notably Dickson et al (2015) [29], in exploring whether the lung microbiome is spatially varied in healthy adults, determined that intrapulmonary sites, when compared to each other, did not contain consistently distinct microbiota, but that intra-subject variation was significantly less than inter-subject variation. We similarly established that whilst sheep lung segment had no significant bearing on the composition of microbiota, sheep identity did have a significant impact. Coupling the latter finding with the sometimes observed high degree of within-lung heterogeneity is conceptually difficult and raises obvious questions regarding both the spatial extent of distinct microbial communities and their longitudinal stability. Indeed, whilst we sampled from disparate lung segments and found heterogeneity in some sheep it is

conceivable that in these animals PSB samples from neighbouring bronchi within the same segment, or even different locations along the same bronchi, might also reflect heterogeneity. Equally we currently have no idea whether lung microbiota 'states' are constant within individual sampling sites over time. Addressing such questions will be fundamental to developing a hypothetical modelling framework for ovine lung microbiota that captures both individual identity and potential within-individual heterogeneity at a given point in time.

Dickson, Erb-Downward and Huffnagle (2014)[32] highlighted in their recent review that the invasive nature of microbiota sampling in the lower respiratory tract has hitherto precluded the gathering of data to assess the extent of temporal and spatial heterogeneity of the lung microbiome in healthy human subjects. We would contend that large animal models offer the facility to probe such relationships and develop experience and methodology that will potentially impact on our ability to understand the relevance of change in the composition of lung microbiota in humans.

Whilst much interest surrounds the relationship between lung microbiota states and diseases such as asthma, COPD and CF, these studies, by their nature, only reflect associations. If such observations are to be usefully extended and the functional significance of lung microbiota established then there is a clear need to develop animal models of lung microbiota states to test mechanistic hypotheses [32].

This model system provided us with the opportunity to determine, in the first instance, whether local lung infection with *P. aeruginosa* would alter lung microbiota, as reflected in PSB samples, in both the direct lung segments and in areas of the lung 'remote' to those segments. Whilst lung infection resulted in consistent lung pathology PSB samples from only two of the direct segments in the placebo group demonstrated a heavy proportional burden of Pseudomonadales. It is worth noting that the airways from which these samples were derived could not be visually differentiated from airways yielding lesser burdens at bronchoscopy (D Collie personal communication). The highly significant positive correlation between qPCR for PA0579 and reads assigned to Pseudomonadales following 16SrRNA sequencing indicates the likelihood that the heavy proportional burden of Pseudomonadales represented the infecting strain, PA0579.

Whilst the failure to demonstrate Pseudomonadalesin PSB samples derived from some subsegmental bronchi serving lung segments with obvious gross pathology and chronic *P. aeruginosa* lung infection presumably reflects the particular pathophysiology underlying these instances, the lack of relationship highlights important caveats in interpreting microbiota changes in our model of lung infection–that PSB microbiota relate only to the precise location wherefrom the sample was derived, and that lung pathology and infection may be highly locally compartmentalised and closely juxtapose airways with minimal evidence of infection.

We also established that local lung infection was not associated with any uniform 'lung-wide' change in lung microbiota sampled from areas distant to that infection. However it was apparent that an increase in the proportion of Pseudomonadales occurred in three of the four remote segments of the two sheep that demonstrated a profound increase in the proportion of Pseudomonadales in their direct lung segments, whereas the sheep that failed to demonstrate an increase in proportion of Pseudomonadales in their direct lung segments also failed to show any appreciable change in the proportion of Pseudomonadales in any of their remote lung segments.

We demonstrate in this ovine model of chronic local lung infection with *P. aeruginosa* that 7 days of once-daily intravenous treatment with CMS had no effect on the burden of *P. aeruginosa* infection in the directly infected lung segments. It is considered that both the nature of the lung pathology and the pharmacodynamics of CMS in sheep following intravenous delivery will have conspired to undermine any therapeutic effect of CMS.

The model protocol, in seeding *P. aeruginosa* in agar beads and delivering these beads to lodge in the distal airways and lung parenchyma evoked a vigorous local inflammatory and immune response that could not be fully resolved. The protective agar matrix, together with abscessation and fibrosis, in representing the body's attempts to limit spread of infection from these chronic nidi of infection, would also potentially hinder access of therapeutics.

CMS has a concentration-dependent effect on Gram negative bacteria. To be effective it must achieve bactericidal peak tissue concentrations in infected lung parenchyma and/or the airway epithelial lining fluid. Whilst we measured neither in this study it is considered unlikely that the dose, mode and frequency of therapy would have generated bactericidal concentrations of CMS in the epithelial lining fluid of the airways. At the inception of this study there was, to the authors' knowledge, no available data concerning the pharmacokinetics of CMS following intravenous administration in sheep, and neither was any specific published guidance available concerning potential toxicity in this species following intravenous delivery. Our choice of dose, mode and frequency of therapy were therefore driven by pragmatic considerations. More recent data does however confirm that CMS is indeed not detected in airway epithelial lining fluid after intravenous dosing in sheep [33]. Further, whilst Boisson et al (2014)[34] did demonstrate an increase in ELF concentration of CMS and colistin after IV administration in critically ill patients, Imberti et al. (2010)[35], following intravenous administration of CMS to adult patients with ventilator-associated pneumonia caused by Gram-negative bacteria, could not measure CMS in BAL fluid. The results of both studies inferred a likely deficiency in therapeutic effect in the lung following intravenous delivery.

In line with these considerations, sheep that were treated with systemic CMS demonstrated a similar spectrum of microbiota change as was seen in the placebo group and there was no evidence of a specific effect of CMS on the proportion of Pseudomonadales detected in PSB samples. However, despite this apparent lack of effect and the unfavourable pharmacokinetic/pharmacodynamic characteristics of intravenous CMS in sheep, therapy did reduce the proportion of Gram negative bacteria (other than Pseudomonadales) and therefore increased the relative proportion of Pseudomonadales to other Gram negative bacteria in these samples. That sub-therapeutic doses can have such an effect may hold relevance when viewed in the context of potentially changing ecological niche characteristics across the whole lung to those more favourable to the survival of *P. aeruginosa*. Indeed, Rogers et al (2014) demonstrated that in non-CF bronchiectasis patients without *P. aeruginosa* airway infection, erythromycin did not significantly reduce exacerbations and promoted displacement of *Haemophilus influenzae* by more macrolide-tolerant pathogens including *P. aeruginosa* [14].

Despite, in some instances, evidence of considerable flux in lung microbiota between baseline and 14 days after establishing chronic local lung infection with *P. aeruginosa*, sheep could not be distinguished on the basis of observed clinical response. Whilst the inference might therefore be construed that local lung microbiota have no functional impact on the pathophysiology of PEs this would be presupposed on the validity of this system to reliably model the pathophysiology and clinical features of PEs in humans. Whilst this presumption has yet to be fully explored in this system, the need to develop novel model systems whereby the pathophysiology of PEs can be investigated means that such studies must remain an imperative.

In conclusion, sheep lung microbiota are dominated in health by bacteria belonging to the orders Bacillales, Actinomycetales and Clostridiales. Whilst chronic local lung infection with *P. aeruginosa* led to increased predominance of Pseudomonadales, such predominance was not uniformly consistent amongst either directly infected or remote lung segments. Treatment of sheep with daily intravenous CMS, whilst failing to overtly influence lung microbiota, did significantly increase the proportion of Pseudomonadales relative to other Gram negative bacteria in infected sheep.

Supporting Information

S1 Fig. Bacterial burden and bronchoalveolar lavage fluid cellularity. Bubble plots depicting the relationship between bronchoalveolar lavage fluid cellular composition and bacterial burden in samples derived at baseline (A. Pre) and after infection, both from areas of the lung subject to direct lung infection with P. aeruginosa (C. Post-direct) and areas remote to those sites (B. Post-remote). Bubbles are coloured according to the sheep to which they relate and the legend further specifies to which treatment group the sheep belong (Treatment_Sheep). The size of each bubble relates to the bacterial burden present in that particular sample and can be gauged through reference to the legend (BAL cfu/ml). Data relating to the log-transformed absolute number of A. alveolar macrophages (Log10 (ABS AMs+1), B. neutrophils (Log10 (ABS Neuts+1), C. lymphocytes Log10 (ABS Lymph+1), D. mast cells Log10 (ABS Mast+1) and E. eosinophils Log10 (ABS Eosin+1) are shown.
(TIFF)

S2 Fig. Bronchoalveolar lavage fluid cytology. Cytospin images representative of baseline (A) and post-infection (B) cytology. Alveolar macrophages were the predominant cell type present in baseline samples. Following the establishment of lung infection the proportion of neutrophils (arrowheads) and lymphocytes (*) increased in the directly infected lung segments.
(TIFF)

Acknowledgments

The authors wish to acknowledge the assistance of Dryden Animal Services in the conduct of this work, and the assistance of Dr Helen Brown in relation to statistical analysis.

Pyrosequencing was carried out by Edinburgh Genomics, The University of Edinburgh. Edinburgh Genomics is partly supported through core grants from NERC (R8/H10/56), MRC (MR/K001744/1) and BBSRC (BB/J004243/1).

Author Contributions

Conceived and designed the experiments: DC GM. Performed the experiments: DC LG SW ET PT GM. Analyzed the data: DC LG SW GM. Contributed reagents/materials/analysis tools: DC LG JG SW PT CD GM. Wrote the paper: DC LG JG SW ET PT CD GM.

References

1. UK Cystic Fibrosis Registry Annual Data Report 2014. Bromley: Cystic Fibrosis Trust, 2015.
2. King PT, Holdsworth SR, Freezer NJ, Villanueva E, Holmes PW. Microbiologic follow-up study in adult bronchiectasis. Respir Med. 2007; 101(8):1633–8. Epub 2007/05/01. doi: 10.1016/j.rmed.2007.03.009 PMID: 17467966.
3. Govan JR, Deretic V. Microbial pathogenesis in cystic fibrosis: mucoid Pseudomonas aeruginosa and Burkholderia cepacia. Microbiol Rev. 1996; 60(3):539–74. PMID: 8840786; PubMed Central PMCID: PMCPMC239456.
4. Sanders DB, Bittner RC, Rosenfeld M, Redding GJ, Goss CH. Pulmonary exacerbations are associated with subsequent FEV1 decline in both adults and children with cystic fibrosis. Pediatr Pulmonol. 2011; 46(4):393–400. doi: 10.1002/ppul.21374 PMID: 20967845.
5. Rogers GB, Zain NMM, Bruce KD, Burr LD, Chen AC, Rivett DW, et al. A Novel Microbiota Stratification System Predicts Future Exacerbations in Bronchiectasis. Annals of the American Thoracic Society. 2014; 11(4):496–503. doi: 10.1513/AnnalsATS.201310-335OC PMID: 24592925
6. Martinez-Solano L, Macia MD, Fajardo A, Oliver A, Martinez JL. Chronic Pseudomonas aeruginosa infection in chronic obstructive pulmonary disease. Clin Infect Dis. 2008; 47(12):1526–33. Epub 2008/11/08. doi: 10.1086/593186 PMID: 18990062.
7. Regelmann WE, Elliott GR, Warwick WJ, Clawson CC. Reduction of sputum Pseudomonas aeruginosa density by antibiotics improves lung function in cystic fibrosis more than do bronchodilators and chest

physiotherapy alone. Am Rev Respir Dis. 1990; 141(4 Pt 1):914–21. Epub 1990/04/01. doi: 10.1164/ajrccm/141.4_Pt_1.914 PMID: 2109558.

8. Burkett A, Vandemheen KL, Giesbrecht-Lewis T, Ramotar K, Ferris W, Chan F, et al. Persistency of Pseudomonas aeruginosa in sputum cultures and clinical outcomes in adult patients with cystic fibrosis. Eur J Clin Microbiol Infect Dis. 2012; 31(7):1603–10. doi: 10.1007/s10096-011-1483-8 PMID: 22108844.
9. Stressmann FA, Rogers GB, Marsh P, Lilley AK, Daniels TW, Carroll MP, et al. Does bacterial density in cystic fibrosis sputum increase prior to pulmonary exacerbation? J Cyst Fibros. 2011; 10(5):357–65. doi: 10.1016/j.jcf.2011.05.002 PMID: 21664196.
10. Reid DW, Latham R, Lamont IL, Camara M, Roddam LF. Molecular analysis of changes in Pseudomonas aeruginosa load during treatment of a pulmonary exacerbation in cystic fibrosis. J Cyst Fibros. 2013. Epub 2013/05/28. doi: 10.1016/j.jcf.2013.03.008 PMID: 23706827.
11. Carmody LA, Zhao J, Schloss PD, Petrosino JF, Murray S, Young VB, et al. Changes in cystic fibrosis airway microbiota at pulmonary exacerbation. Ann Am Thorac Soc. 2013; 10(3):179–87. doi: 10.1513/AnnalsATS.201211-107OC PMID: 23802813.
12. Price KE, Hampton TH, Gifford AH, Dolben EL, Hogan DA, Morrison HG, et al. Unique microbial communities persist in individual cystic fibrosis patients throughout a clinical exacerbation. Microbiome. 2013; 1(1):27. Epub 2014/01/24. doi: 10.1186/2049-2618-1-27 PMID: 24451123; PubMed Central PMCID: PMCPmc3971630.
13. Daniels TW, Rogers GB, Stressmann FA, van der Gast CJ, Bruce KD, Jones GR, et al. Impact of antibiotic treatment for pulmonary exacerbations on bacterial diversity in cystic fibrosis. J Cyst Fibros. 2013; 12(1):22–8. doi: 10.1016/j.jcf.2012.05.008 PMID: 22717533.
14. Rogers GB, Bruce KD, Martin ML, Burr LD, Serisier DJ. The effect of long-term macrolide treatment on respiratory microbiota composition in non-cystic fibrosis bronchiectasis: an analysis from the randomised, double-blind, placebo-controlled BLESS trial. Lancet Respir Med. 2014; 2(12):988–96. Epub 2014/12/03. doi: 10.1016/s2213-2600(14)70213-9 PMID: 25458200.
15. Nicholas B, Djukanovic R. Induced sputum: a window to lung pathology. Biochem Soc Trans. 37. England2009. p. 868–72. doi: 10.1042/BST0370868 PMID: 19614609
16. Collie D, Govan J, Wright S, Thornton E, Tennant P, Smith S, et al. A lung segmental model of chronic Pseudomonas infection in sheep. PLoS One. 2013; 8(7):e67677. doi: 10.1371/journal.pone.0067677 PMID: 23874438; PubMed Central PMCID: PMCPMC3706528.
17. Li J, Milne RW, Nation RL, Turnidge JD, Coulthard K. Stability of colistin and colistin methanesulfonate in aqueous media and plasma as determined by high-performance liquid chromatography. Antimicrob Agents Chemother. 2003; 47(4):1364–70. Epub 2003/03/26. PMID: 12654671; PubMed Central PMCID: PMCPMC152538.
18. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Meth. 2012; 9(7):671–5.
19. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med. 2011; 184(8):957–63. Epub 2011/06/18. doi: 10.1164/rccm.201104-0655OC PMID: 21680950; PubMed Central PMCID: PMCPmc3208663.
20. Govan JRW, Fyfe JAM. Mucoid *Pseudomonas aeruginosa* and cystic fibrosis: resistance of the mucoid form to carbenicillin, flucloxacillin and tobramycin and the isolation of mucoid variants *in vitro*. Journal of Antimicrobial Chemotherapy. 1978; 4(3):233–40. doi: 10.1093/jac/4.3.233 PMID: 97259
21. Goddard AF, Staudinger BJ, Dowd SE, Joshi-Datar A, Wolcott RD, Aitken ML, et al. Direct sampling of cystic fibrosis lungs indicates that DNA-based analyses of upper-airway specimens can misrepresent lung microbiota. Proceedings of the National Academy of Sciences. 2012; 109(34):13769–74. doi: 10.1073/pnas.1107435109
22. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq illumina sequencing platform. Applied and Environmental Microbiology. 2013; 79(17):5112–20. doi: 10.1128/aem.01043-13 PMID: 23793624
23. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 2011;17(1):10–2.
24. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology. 2009; 75(23):7537–41. doi: 10.1128/aem.01541-09 PMID: 19801464

25. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011; 27(16):2194–200. doi: 10.1093/bioinformatics/btr381 PMID: 21700674

26. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Applied and Environmental Microbiology. 2006; 72(7):5069–72. doi: 10.1128/aem.03006-05 PMID: 16820507

27. Le Gall F, Le Berre R, Rosec S, Hardy J, Gouriou S, Boisrame-Gastrin S, et al. Proposal of a quantitative PCR-based protocol for an optimal Pseudomonas aeruginosa detection in patients with cystic fibrosis. BMC Microbiol. 13. England2013. p. 143. doi: 10.1186/1471-2180-13-143 PMID: 24088260

28. Segal LN, Alekseyenko AV, Clemente JC, Kulkarni R, Wu B, Chen H, et al. Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. Microbiome. 2013; 1 (1):19. doi: 10.1186/2049-2618-1-19 PMID: 24450871; PubMed Central PMCID: PMCPMC3971609.

29. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Beck JM, Huffnagle GB, et al. Spatial Variation in the Healthy Human Lung Microbiome and the Adapted Island Model of Lung Biogeography. Ann Am Thorac Soc. 2015. Epub 2015/03/25. doi: 10.1513/AnnalsATS.201501-029OC PMID: 25803243.

30. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. PLoS One. 2010; 5(1):e8578. doi: 10.1371/journal.pone.0008578 PMID: 20052417; PubMed Central PMCID: PMCPMC2798952.

31. Bekele AZ, Koike S, Kobayashi Y. Genetic diversity and diet specificity of ruminal Prevotella revealed by 16S rRNA gene-based analysis. FEMS Microbiol Lett. 305. England2010. p. 49–57. doi: 10.1111/j.1574-6968.2010.01911.x PMID: 20158525

32. Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. Expert Rev Respir Med. 2013; 7(3):245–57. Epub 2013/06/06. doi: 10.1586/ers.13.24 PMID: 23734647; PubMed Central PMCID: PMCPmc4007100.

33. Nguyen T, Landersdorfer C, Lieu L, Bischof R, Meeusen E, McIntosh M, editors. Abstract 59. Targeted delivery of antimicrobials to the lung: sheep as a clinically relevant model to assess the pharmacokinetics of the colistin prodrug colistin methansulfonate after pulmonary and intravenous administration. Journal of Aerosol Medicine and Pulmonary Drug Delivery; 2013; Edinburgh, Scotland: Mary Ann Liebert, Inc 140 Huguenot Street, 3rd FL, New Rochelle, NY 10801 USA.

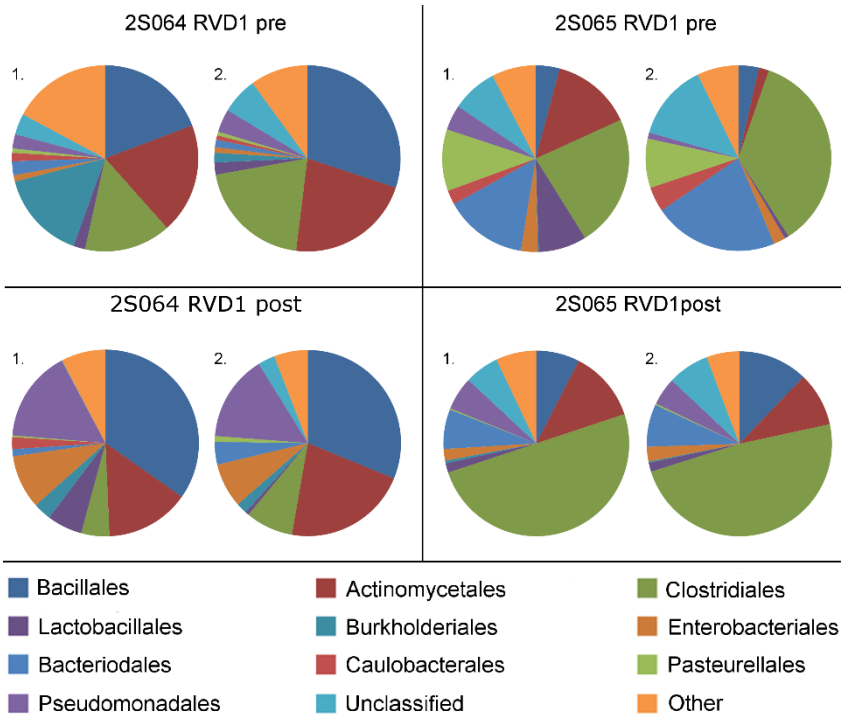
34. Boisson M, Jacobs M, Gregoire N, Gobin P, Marchand S, Couet W, et al. Comparison of intrapulmonary and systemic pharmacokinetics of colistin methanesulfonate (CMS) and colistin after aerosol delivery and intravenous administration of CMS in critically ill patients. Antimicrob Agents Chemother. 58. United States: American Society for Microbiology. All Rights Reserved.; 2014. p. 7331–9. doi: 10.1128/AAC.03510-14 PMID: 25267660

35. Imberti R, Cusato M, Villani P, Carnevale L, Iotti GA, Langer M, et al. Steady-state pharmacokinetics and BAL concentration of colistin in critically Ill patients after IV colistin methanesulfonate administration. Chest. 138. United States2010. p. 1333–9. doi: 10.1378/chest.10-0463 PMID: 20558557

Appendix 2: Supplemental figures and tables for Chapter 3

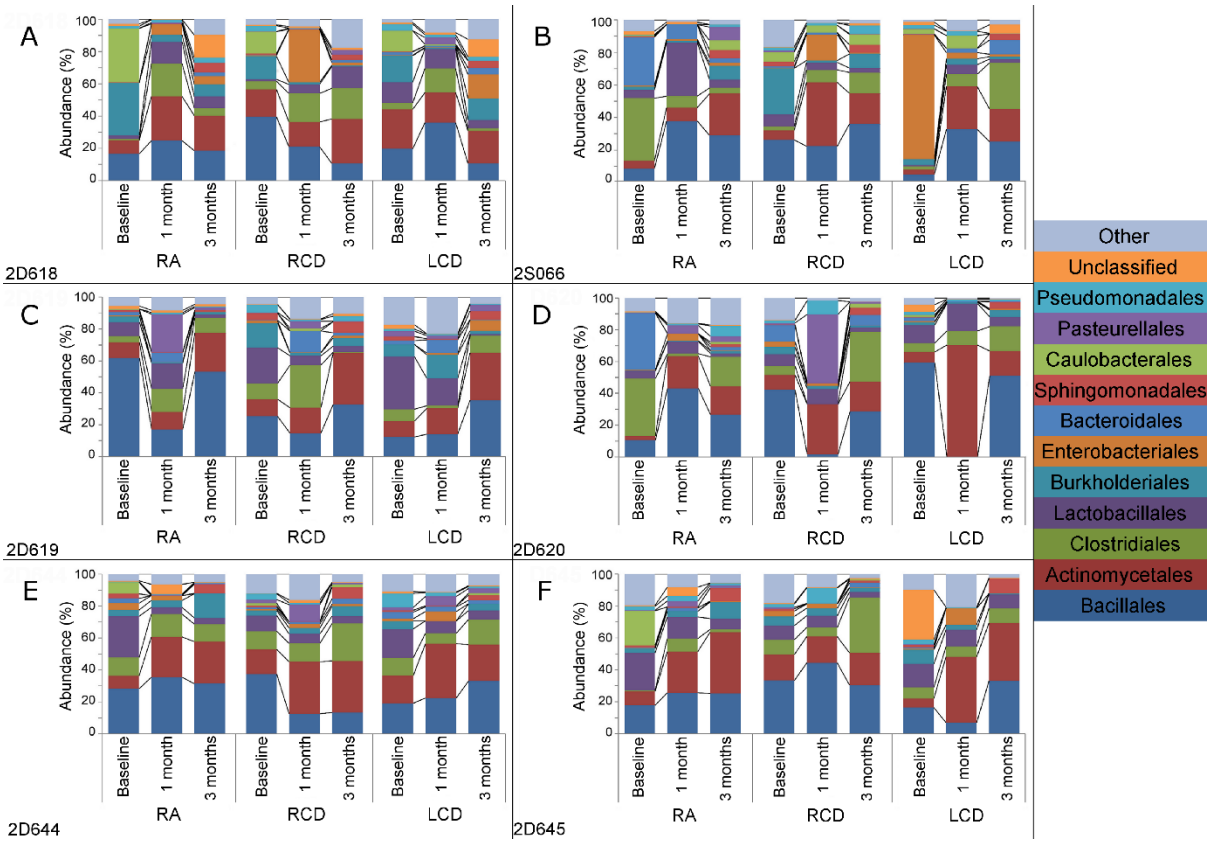
Supplemental material

Figure S1



Comparison of bacterial communities found in lung brushings from a previous study (1). Protected specimen brushings were taken in two sheep (2S064 and 2S065) from the first ventral diaphragmatic segment of the right caudal diaphragmatic lobe (RVD1). Samples were taken pre and post injection with either colisitin (2S065) or saline (2S064). The bacterial DNA in these samples was amplified twice; once in May 2013 (1.) and once in September 2013 (2.). These amplicons were then sequenced and bioinformatically processed separately from one another. Despite some small differences, amplicons produced from the same sample are highly similar to one another.

Figure S2



The bacterial orders found at three separate lung segments (RA: right apical, RCD: right caudal diaphragmatic and LCD: left caudal diaphragmatic) in six sheep (A: 2D618, B: 2S066, C: 2D619, D: 2D620, E: 2D644, F: 2D645) at three time-points: baseline (day 0), one month and three months.

Table S1: Dates of bronchial brushing samplings for six sheep at three time-points.

Time-point	Sheep	Date
Baseline (d0)	2D618	18/11/2013
	2S066	18/11/2013
	2D619	18/11/2013
	2D620	20/11/2013
	2D644	20/11/2013
	2D645	20/11/2013
One Month	2D618	16/12/2013
	2S066	16/12/2013
	2D619	16/12/2013
	2D620	17/12/2013
	2D644	17/12/2013
	2D645	17/12/2013
Three Months	2D618	17/02/2014
	2S066	17/02/2014
	2D619	17/02/2014
	2D620	18/02/2014
	2D644	18/02/2014
	2D645	18/02/2014

Table S2: Barcoded primer sequences used during the second round of PCR

Primer	Sequences
104f501	5'~AATGATACGGCGACCACCGAGATCTACACTATAGCCTACACTCT TTCCCTACACGACGCTCTTCCGATCTGGCGVACGGGTGAGTAA~3'
104f502	5'~AATGATACGGCGACCACCGAGATCTACACATAGAGGCACACTCT TTCCCTACACGACGCTCTTCCGATCTNGGCGVACGGGTGAGTAA~3'
104f503	5'~AATGATACGGCGACCACCGAGATCTACACCCTATCCTACACTCTT TCCCTACACGACGCTCTTCCGATCTNNGGCGVACGGGTGAGTAA ~3'
104f504	5'~AATGATACGGCGACCACCGAGATCTACACGGCTCTGAACACTCT TTCCCTACACGACGCTCTTCCGATCTNNNGGCGVACGGGTGAGTAA~ 3'
104f505	5'~AATGATACGGCGACCACCGAGATCTACACAGGCGAAGACACTCT TTCCCTACACGACGCTCTTCCGATCTGGCGVACGGGTGAGTAA~3'
104f506	5'~AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTT TCCCTACACGACGCTCTTCCGATCTGGCGVACGGGTGAGTAA~3'
104f507	5'~AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCT TTCCCTACACGACGCTCTTCCGATCTGGCGVACGGGTGAGTAA~3'
104f508	5'~AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCT TTCCCTACACGACGCTCTTCCGATCTGGCGVACGGGTGAGTAA~3'
519r701	5'~CAAGCAGAAGACGGCATACGAGATATTACTCGGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTNNNGTNTTACNGCGGCKGCTG~3'
519r702	5'~CAAGCAGAAGACGGCATACGAGATTCCGGAGAGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTNNGTNTTACNGCGGCKGCTG~3'

519r703	5’~CAAGCAGAAGACGGCATACGAGATCGCTCATTGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTNGTNTTACNGCGGCKGCTG~3’
519r704	5’~CAAGCAGAAGACGGCATACGAGATGAGATTCCGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTGTNTTACNGCGGCKGCTG~3’
519r705	5’~CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTGTNTTACNGCGGCKGCTG~3’
519r706	5’~CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTGTNTTACNGCGGCKGCTG~3’
519r707	5’~CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTGTNTTACNGCGGCKGCTG~3’
519r708	5’~CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTT CAGACGTGTGCTCTTCCGATCTGTNTTACNGCGGCKGCTG~3’

Table S3: OTUs found to be significantly different between lung samples taken from sheep at the baseline and one month time-points (Metastats: P = 0.001)

OTU phylotype	Relative abundance (%) mean ± SEM *	
	Baseline	One month
Bacteroides	0.69 ± 0.51	0 ± 0
Bradyrhizobium	0.19 ± 0.069	0.0051 ± 0.0051
Burkholderia	3.34 ± 1.69	0 ± 0
Burkholderia graminis	0.28 ± 0.15	0 ± 0
Comamonas	0.028 ± 0.0001	0 ± 0
Corynebacterium	0.72 ± 0.19	11.05 ± 3.54
Corynebacterium kroppenstedtii	0.0025 ± 0.0025	2.97 ± 0.57
Gallicola	0 ± 0	0.14 ± 0.095
Granulicatella	3.85 ±0.62	0.82 ± 0.42
MBA08	0 ± 0	0.12 ± 0.12
Methylobacteriaceae	0.22 ± 0.093	0 ± 0
Methylobacterium organophilum	0.12 ± 0.054	0 ± 0
Mycobacterium llatzerense	0.015 ± 0.009	2.47 ± 0.97
Roseburia	0.10 ± 0.081	0 ± 0
Ruminococcus	0.10 ± 0.081	0 ± 0

Sphingomonas yabuuchiae	0.78 ± 0.22	0 ± 0
Staphylococcus haemolyticus	0 ± 0	0.63 ± 0.63
Streptococcus infantis	1.51 ± 0.26	0.025 ± 0.02
Tannerella	0.12 ± 0.12	0 ± 0
Tepidimonas	0 ± 0	0.25 ± 0.25
Tetragenococcus	0.11 ± 0.11	0 ± 0

* OTUs are not included where they were less than 0.1% abundant at both time-points

Table S4: OTUs found to be significantly different between lung samples taken from sheep at the one month and three month time-points (Metastats: P = 0.001)

OTU phylotype	Relative abundance (%) mean ± SEM*	
	One month	Three months
Actinobacillus parahaemolyticus	0 ± 0	0.12 ± 0.11
Bacteroidales	0 ± 0	0.40 ± 0.39
Blastomonas	0.0051 ± 0.0051	1.5 ± 0.18
Blautia	0.20 ± 0.20	0 ± 0
Caulobacter	0 ± 0	0.21 ± 0.16
Cohnella	0 ± 0	0.11 ± 0.11
Comamonas	0 ± 0	0.76 ± 0.21

Helcococcus	0.23 ± 0.16	0 ± 0
MBA08	0.12 0.12	0 ± 0
Mycobacterium	0.048 ± 0.046	1.9 ± 0.46
Mycobacterium llatzerense	2.47 ± 0.98	0 ± 0
Myroides odoratimimus	0 ± 0	0.16 ± 0.16
Novosphingobium	0.018 ± 0.013	0.32 ± 0.096
Oligella	0 ± 0	0.11 ± 0.11
Prevotella copri	0 ± 0	0.18 ± 0.18
Prevotella tannerae	0 ± 0	0.12 ± 0.12
Rubricoccus	0.35 ± 0.25	0 ± 0
Schlegelella	0 ± 0	0.11 ± 0.11
Sneathia	0.20 ± 0.20	0 ± 0
Sphingobium yanoikuyae	0 ± 0	2.70 ± 0.31
Tepidimonas	0.25 ± 0.25	0 ± 0

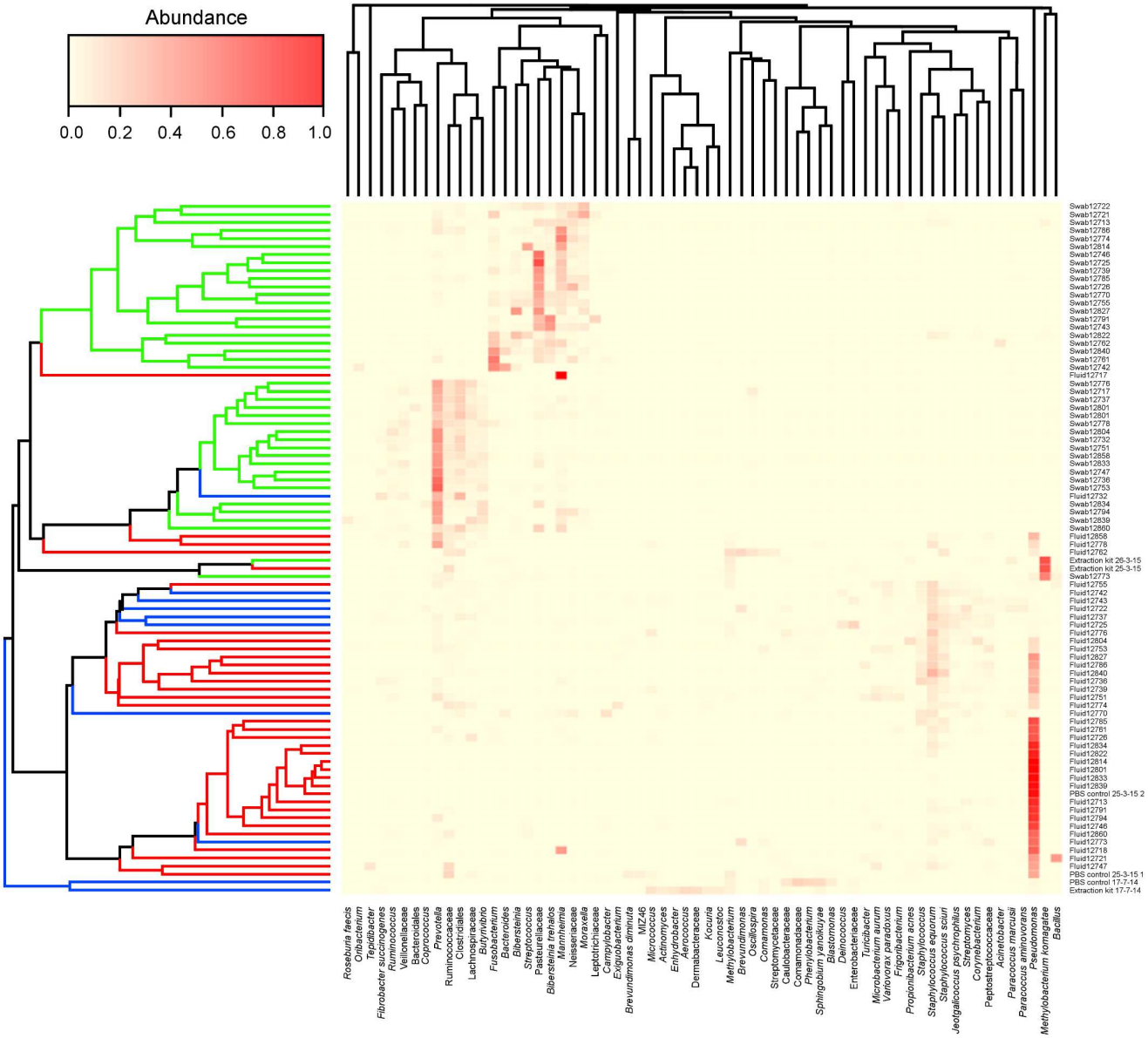
* OTUs are not included where they were less than 0.1% abundant at both time-points

References

1. **Collie D, Glendinning L, Govan J, Wright S, Thornton E, Tennant P, Doherty C, McLachlan G.** 2015. Lung microbiota changes associated with chronic *Pseudomonas*

aeruginosa lung infection and the impact of intravenous colistimethate. PloS ONE **10:**e0142097. DOI: 10.1371/journal.pone.0142097

Appendix 3: Figure S1 for Chapter 4



Heatmap of OTUs found in lamb lung fluids, oropharyngeal swabs, PBS and extraction kit reagent only controls. OTUs were included when they were > 5% abundant in at least one sample. Samples whose DNA was extracted on separate days are indicated by the following colours on the hierarchical clustering dendrogram: 17th July 2014 = blue; 25th March 2015 = red; 26th March 2015 = green.

Appendix 4: Table S1 for Chapter 4

OTUs responsible for partitioning of lamb oropharyngeal swabs into two groups (using Laplace value).

Taxonomy	Indicator groups	P value
Bacteroides	Partition 1	<0.001
Bibersteinia trehalosi	Partition 1	<0.001
Campylobacter rectus	Partition 1	<0.001
Leptotrichia	Partition 1	<0.001
Moraxella	Partition 1	<0.001
Porphyromonas	Partition 1	<0.001
Streptococcus	Partition 1	<0.001
Streptococcus minor	Partition 1	<0.001
Kingella	Partition 1	0.003
Moraxella ovis	Partition 1	0.003
Pasteurellaceae	Partition 1	0.003
Comamonadaceae	Partition 1	0.004
Cardiobacteriaceae	Partition 1	0.005
Corynebacterium	Partition 1	0.005
Lactobacillales	Partition 1	0.005
Tannerella	Partition 1	0.005
BD1-5	Partition 1	0.006
Bibersteinia	Partition 1	0.006
Fusobacterium	Partition 1	0.007

Mannheimia	Partition 1	0.007
Streptococcus agalactiae	Partition 1	0.007
Actinomyces hyovaginalis	Partition 1	0.008
Aggregatibacter	Partition 1	0.008
Caulobacteraceae	Partition 1	0.008
Methylobacterium komagatae	Partition 1	0.008
Neisseriaceae	Partition 1	0.008
Conchiformibius kuhniae	Partition 1	0.009
Variovorax paradoxus	Partition 1	0.009
Dietzia	Partition 1	0.01
Hylemonella	Partition 1	0.01
Streptococcaceae	Partition 1	0.01
Lautropia	Partition 1	0.011
Frigoribacterium	Partition 1	0.014
Micrococcaceae	Partition 1	0.014
Peptostreptococcaceae	Partition 1	0.014
Actinomyces	Partition 1	0.016
Streptobacillus	Partition 1	0.016
Staphylococcus	Partition 1	0.019
Staphylococcus sciuri	Partition 1	0.024
Campylobacter	Partition 1	0.025
Peptostreptococcus anaerobius	Partition 1	0.026
Propionibacteriaceae	Partition 1	0.026
Brachybacterium	Partition 1	0.027
Erysipelotrichaceae	Partition 1	0.029

Fusobacteriaceae	Partition 1	0.029
Parvimonas	Partition 1	0.029
Streptococcus equi	Partition 1	0.032
Facklamia	Partition 1	0.037
Microvirgula	Partition 1	0.04
Clavibacter michiganensis	Partition 1	0.043
Sphingomonas echinoides	Partition 1	0.043
Staphylococcus equorum	Partition 1	0.045
Acinetobacter venetianus	Partition 1	0.046
Acinetobacter lwoffii	Partition 1	0.047
Arthrobacter	Partition 1	0.047
Bacteroidales	Partition 2	<0.001
Blautia	Partition 2	<0.001
Butyrivibrio	Partition 2	<0.001
Carboxydocellaceae	Partition 2	<0.001
CF231	Partition 2	<0.001
Clostridiales	Partition 2	<0.001
Clostridium	Partition 2	<0.001
Coprococcus	Partition 2	<0.001
Coriobacteriaceae	Partition 2	<0.001
Desulfovibrio D168	Partition 2	<0.001
Eggerthella	Partition 2	<0.001
Faecalibacterium	Partition 2	<0.001
Fibrobacter succinogenes	Partition 2	<0.001
Lachnospiraceae	Partition 2	<0.001

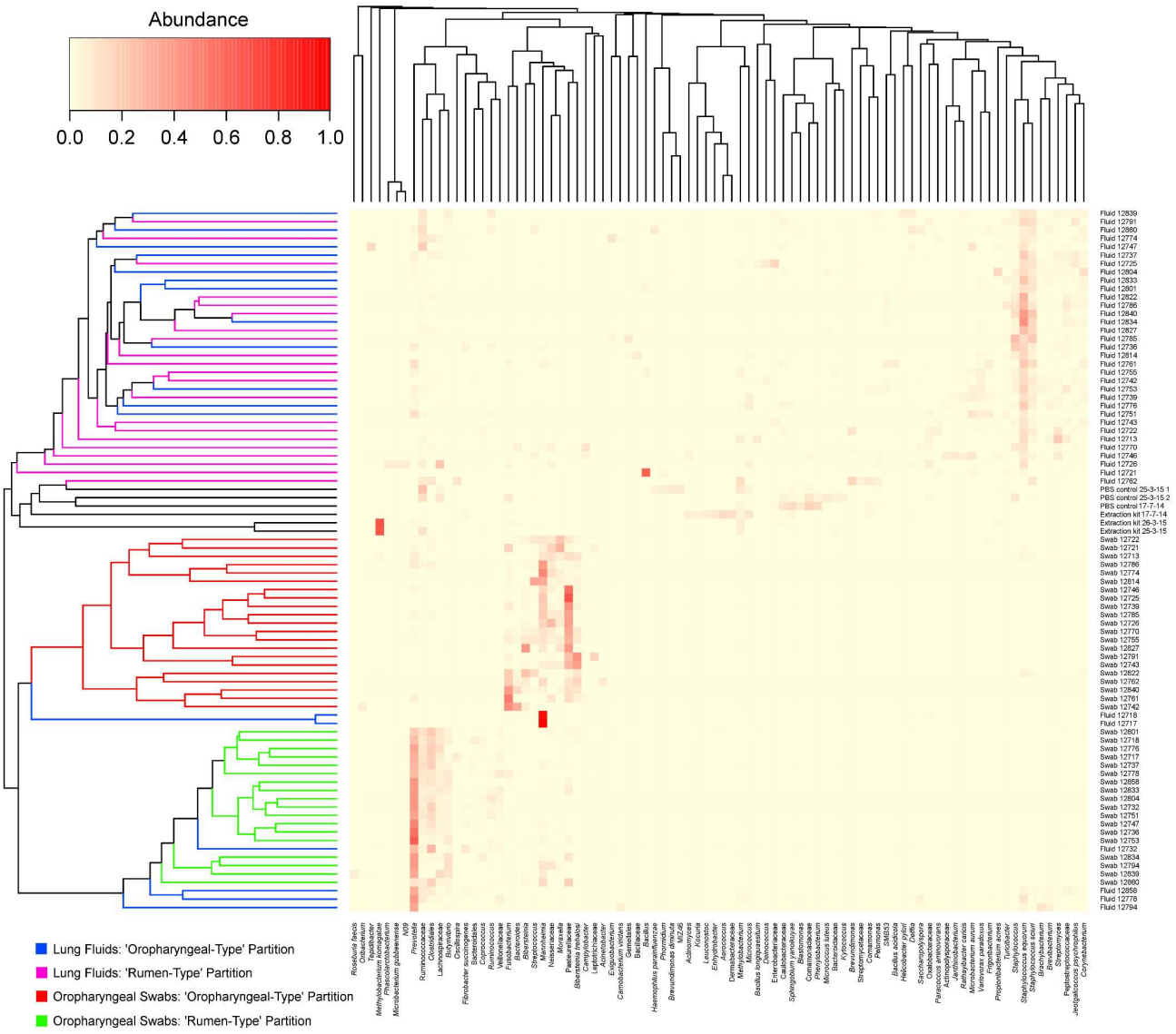
Mogibacteriaceae	Partition 2	<0.001
Mogibacterium	Partition 2	<0.001
Moryella	Partition 2	<0.001
Moryella indoligenes	Partition 2	<0.001
Oscillospira	Partition 2	<0.001
Oscillospira guilliermondii	Partition 2	<0.001
p-75-a5	Partition 2	<0.001
Paraprevotellaceae	Partition 2	<0.001
Persicobacter	Partition 2	<0.001
Prevotella	Partition 2	<0.001
Prevotellaceae	Partition 2	<0.001
Prevotella ruminicola	Partition 2	<0.001
Pseudobutyrivibrio	Partition 2	<0.001
Pyramidobacter	Partition 2	<0.001
RFN20	Partition 2	<0.001
Ruminococcaceae	Partition 2	<0.001
Ruminococcus flavefaciens	Partition 2	<0.001
S24-7	Partition 2	<0.001
SHD-231	Partition 2	<0.001
Shuttleworthia	Partition 2	<0.001
Succiniclasticum	Partition 2	<0.001
TTA_B6	Partition 2	<0.001
Veillonellaceae	Partition 2	<0.001
YRC22	Partition 2	<0.001
YS2	Partition 2	<0.001

Desulfovibrio	Partition 2	0.002
Ruminococcus	Partition 2	0.004
Fibrobacteraceae	Partition 2	0.006
Treponema	Partition 2	0.006
RF39	Partition 2	0.007
Chloroherpetales	Partition 2	0.009
Thermogemmatisporales	Partition 2	0.012
Saprospirae	Partition 2	0.013
L7A_E11	Partition 2	0.014
SSW63Au	Partition 2	0.015
Anaerofustis	Partition 2	0.017
Christensenellaceae	Partition 2	0.018
Elizabethkingia	Partition 2	0.022
Anaerostipes	Partition 2	0.023
Bosea	Partition 2	0.024
Spirochaetaceae	Partition 2	0.024
Kurthia	Partition 2	0.029
BS11	Partition 2	0.031
Methylacidiphilae	Partition 2	0.033
Anaeroplasma	Partition 2	0.034
Rhodothermales	Partition 2	0.036
BF311	Partition 2	0.044
Deferribacterales	Partition 2	0.045
Desulfovibrionaceae	Partition 2	0.048

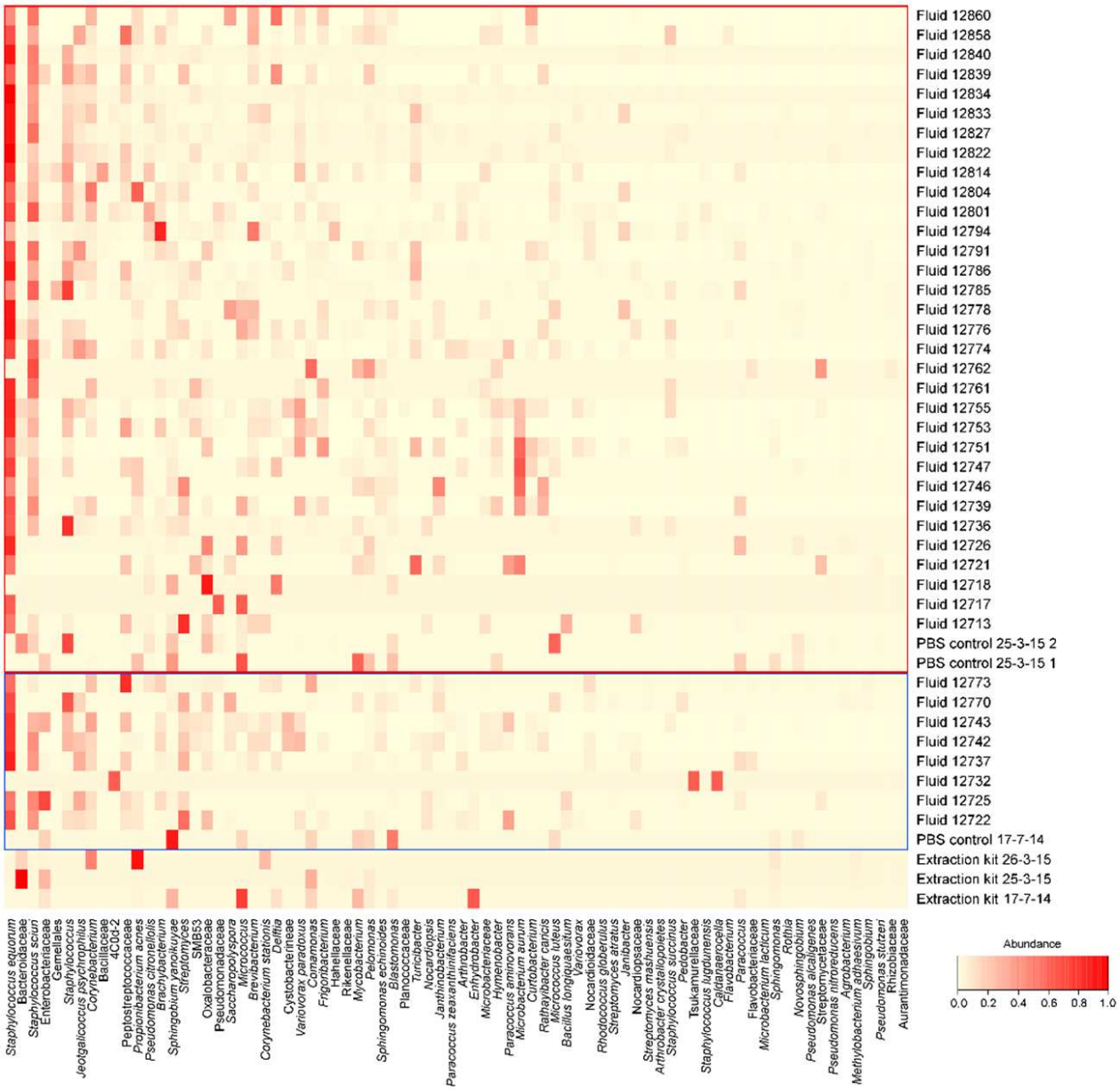
Selenomonas ruminantium	Partition 2	0.048

The OTUs sorted into partition 1 are more oropharyngeal-like whereas those sorted into partition 2 are more rumen-like. The OTUs which were significantly indicative of either partition 1 or 2 were calculated using the indicator command within mothur.

Appendix 5: Large version of Figure 4.2



Appendix 6: Large version of Figure 4.3



Appendix 7: Large version of Chapter 5: Figure 1